Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity

Pao, Kuan-Chuan; Wood, Nicola T.; Knebel, Axel; Rafie, Karim; Stanley, Mathew; Mabbitt, Peter D.

Published in:
Nature

DOI:
10.1038/s41586-018-0026-1

Publication date:
2018

Document Version
Peer reviewed version

Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity

Kuan-Chuan Pao[1], Nicola T. Wood[1], Axel Knebel[1], Karim Rafie[2], Mathew Stanley[2], Peter D. Mabbitt[1], Ramasubramanian Sundaramoorthy[2], Kay Hofmann[3], Daan van Aalten[2] and Satpal Virdee[1]*

[1]MRC Protein Phosphorylation and Ubiquitylation Unit, University of Dundee, Scotland, UK, DD1 5EH

[2]Division of Gene Regulation and Expression, School of Life Sciences, University of Dundee, Scotland, UK, DD1 5EH

[3]Institute for Genetics, University of Cologne, Germany

* Corresponding author

Email: s.s.virdee@dundee.ac.uk

Tel: +44 (0)1382 388738

Fax: +44 (0)1382 388500

Ubiquitination is initiated by ubiquitin (Ub) transfer from an E1 activating enzyme (E1) to an E2 conjugating enzyme (E2) producing a covalently linked intermediate (E2~Ub)[1]. E3 ligases (E3s) of the Really Interesting New Gene (RING) class recruit E2~Ub via their RING domain and mediate direct transfer of Ub to substrates[2]. In contrast, Homologous to E6-AP Carboxy Terminus (HECT) E3s undergo a catalytic cysteine-dependent transthiolation reaction with E2~Ub forming a covalent E3~Ub intermediate[3,4]. Additionally, RING-between-RING (RBR) E3s have a canonical RING domain that is linked to an ancillary domain. This contains a catalytic cysteine enabling a hybrid RING/HECT mechanism[5]. Ubiquitination is typically considered a posttranslational modification of lysine residues as human E3s endowed with non-lysine activity remain to be discovered. Herein, we carry out activity-based protein profiling of HECT/RBR-like E3s and uncover the neuron-associated E3 MYCBP2/Phr1 as a novel class of RING-linked E3 with esterification activity and intrinsic selectivity for threonine over serine. MYCBP2 contains two essential catalytic cysteine residues which relay Ub to substrate via thioester intermediates. Crystallographic characterization of this new class of E3 ligase, which we designate as an RING-Cys-Relay (RCR), reveals insights into its mechanism and threonine selectivity. These findings implicate cellular regulation of higher eukaryotes by non-lysine ubiquitination and unappreciated mechanistic diversity of E3 enzymes.

We prepared biotinylated variants of our recently developed activity-based probes (ABPs)[6] which profile the hallmark transthiolation activity of HECT/RBR E3s (**Figure 1a and Extended Figure 1**)[7]. Interfacing the ABP technology with mass spectrometry enabled parallelized profiling of E3 activity in neuroblastoma SH-SY5Y cell extracts[8,9] **(Figure 1b and Supplementary Information)**. E3s were filtered using criteria ensuring signals for at least a

subset of detected E3s correlated with E3 activity and/or abundance **(Extended Data Figure 2)**. Profiling of ~80% of the ~50 known HECT/RBR E3s was achieved but unexpectedly, 33 RING E3s, devoid of HECT or RBR ancillary domains, were also enriched **(Figure 1c-e)**. To explore the possibility that hitherto undiscovered RING-linked E3s were being labelled we focused on MYCBP2/Phr1 (Myc-binding protein 2; PAM/Highwire/Rpm-1) (**Figure 1c**). MYCBP2 is a large 0.5 MDa neuron-associated protein which contains a C-terminal RING domain (**Figure 2a**) and is involved in a range of cellular processes including regulation of nervous system development and axon degeneration[10-12].

A recombinant C-terminal version of MYCBP2 encompassing the RING domain (residues 4378-4640; $MYCBP2_{cat}$; **Figure 2a**) and an uncharacterized C-terminal cysteine-rich region underwent robust ABP labelling with an efficiency comparable to that of E3s known to demonstrate transthiolation activity[7,13] **(Extended Data Figure 3a)**. To map the putative catalytic cysteine we used a combination of ABP-based profiling and ABP-crosslinking MS[7] **(Figure 2b, Extended Data Figure 3b, c, full gels and blots in Supplementary Figure 1)**. Data were in support of C4520 being a putative catalytic residue. We next assayed wild type (WT) E3 activity but were unable to detect autoubiquitination or free Ub chain formation. However, we observed rapid E3-dependent discharge of Ub from E2~Ub suggestive of the presence of an unknown small molecule nucleophilic acceptor (**Figure 2c**). Liquid chromatography-mass spectrometry (LC-MS) analysis revealed that Ub was being quantitatively converted into two species with masses corresponding to condensation products with *tris*(hydroxymethyl)aminomethane (Tris) and glycerol (8639 and 8668 Da), both of which were present in our assay buffer at routinely employed concentrations of 50 mM and ~65 mM, respectively. Due to the common hydroxy functionality within these

nucleophiles, MYCBP2 appeared to have esterification activity (**Extended Data Figure 4a, b)**. Activity was found to be dependent on C4520, consistent with it forming a thioester-linked E3~Ub intermediate[4,5] **(Figure 2c)**.

Unexpectedly, a MYCBP2 C4572S mutant retained activity yet formed a discrete mono Ub adduct that was resistant to thiolysis but reversible after base treatment (**Figure 2c, d and Extended Data Figure 4c)**[5]. A possibility was the mutated S4572 residue was contributing to catalysis via formation of an additional and less transient (oxy)ester-linked intermediate between Ub and S4572 that retained the ability to modify substrate**.** We hypothesized that C4520 and C4572 were both catalytic residues that functioned in tandem by relaying Ub from one cysteine to the other through an intramolecular transthiolation reaction. To test this relay mechanism, we carried out gel-based thioester/ester trapping assays[14] (**Figure 2e and Extended Data Figure 4d)** and observed a thiol-sensitive Ub adduct on WT MYCBP2$_{cat}$ which was not observed with the C4520S mutant (**Figure 2e**). Consistent with earlier experiments (**Figure 2c**), C4572S underwent adduct formation that was thiol-resistant but base labile (**Figure 2e**). Thus, if a transient thioester intermediate was being formed between Ub and C4520 an unreactive C4572A mutant should stabilize it. Indeed, thiol-sensitive adduct formation was increased on a C4572A mutant relative to that of the wild type and was presumably linked via a thioester bond **(Figure 2e**). Adduct formation was not detected with a C4520A/C4572S double mutant in support of it being linked to the C4520 residue (**Figure 2e**). In the absence of direct demonstration of Cys-to-Cys Ub transfer we cannot formally exclude other possibilities. However, the existing data are consistent with the essential cysteines functioning in a relay mechanism. Mutational analysis and size exclusion chromatography-multi angle light scattering (SEC-MALS) of MYCBP2$_{cat}$ (**Extended**

Data Figure 4e, f) suggests the proposed relay mechanism requires both essential cysteines be in the same molecule, consistent with an intramolecular *cis* relay mechanism.

In light of the observed esterification activity we attempted to identify the amino acid substrate of MYCBP2 by screening a panel of amino acids[5] and found that discharge activity was strikingly enhanced towards threonine. Product formation was dependent on C4520 (**Figure 3a, b** and **Extended Data Figure 5a-d**). MS-based quantification indicated ~10-fold selectivity for threonine over serine **(Extended Data Figure 5e)**. Although a low level of lysine modification was observed, this was independent of $MYCBP2_{cat}$[5]. Threonine selectivity (3-fold) was also maintained in a peptide context (**Figure 3c** and **Extended Data Figure 5f-h**). Furthermore, basal ubiquitination of a lysine peptide was partially inhibited in the presence of $MYCBP2_{cat}$ underscoring its lack of lysine activity (**Figure 3c**). Taken together, our experiments revealed that MYCBP2 is a novel class of E3 enzyme that operates via two essential cysteines, promotes Ub modification of hydroxyl groups, and esterifies threonine with Ub with selectivity over serine. As MYCBP2 uses a novel mechanism we termed it a RING-Cys-Relay (RCR) E3. We benchmarked the catalytic efficiency of MYCBP2 threonine esterification activity and found it to fall between that of well-characterized HECT and RBR E3 lysine aminolysis activity[5,15] (**Extended Data Figure 5i-k)**. E2 mutational analysis[5,16-19], was in further support of $MYCBP2_{cat}$ being a novel class of E3 (**Extended Data Figure 6a, b and Methods).** To ascertain functional E2 partners, 17 E2s were tested but only UBE2D1, UBE2D3 and UBE2E1 demonstrated robust activity (**Extended Data Figure 6c)**.

MYCBP2 promotes Wallerian axon degeneration through destabilization of Nicotinamide Mononucleotide Adenyltransferase (NMNAT2)[20]. We next tested whether $MYCBP2_{cat}$ can ubiquitinate NMNAT2 by esterification *in vitro* (**Figure 3d**). Despite containing 13 lysine residues, NMNAT2 underwent hydroxide labile but thiol resistant ubiquitination demonstrating that MYCBP2 can target hydroxy residues within one of its putative substrates[20]. Cellular substrate recognition is mediated by a Skp1/Fbox45 substrate receptor co-complex that binds to a site ~1940 residues N-terminal to the $MYCBP2_{cat}$ region (**Figure 2a**)[21]. NMNAT2 also undergoes palmitoylation and rapid axonal transport[22] making reconstitution and cellular study of its ubiquitination extremely challenging. However, to establish whether $MYCBP2_{cat}$ retains non-lysine activity in cells we looked at its autoubiquitination after transient transfection into human embryonic kidney 293 cells. Base-labile (but thiol-resistant) ubiquitination was observed that was dependent on C4520 (**Figure 3e**). This demonstrates that MYCBP2 can retain specificity for hydroxy amino acids in cells and that this activity remains dependent on the upstream catalytic residue we implicate with a Ub relay mechanism.

To further validate the RING-Cys-relay model and the serine/threonine activity we crystallized $MYCBP2_{cat}$ (residues 4378-4640) and solved a crystal structure to a resolution 1.75 Å (**Extended Data Table 1 and Extended Data Figure 7a-c**). Residues 4388-4441 at the N-terminus correspond to the predicted cross-brace C3H2C3 RING domain (**Figure 4a and Extended Data Figure 7d**). Following the RING domain is a long α-helix (4447-4474) that leads into small helix-turn-helix motif (residues 4475-4500) (**Figure 4a and Extended Data Figure 7e**) and further C-terminal is a structurally unprecedented globular domain that binds four Zn ions (residues 4501-4638) (**Figure 4b and Extended Data 7 f, g**). Since this domain

also contains the two essential catalytic residues we term it the Tandem Cysteine (TC) domain. Between the βA2 strand and helix 3_{10}A is an unstructured region (4519-4526) which projects out to the side of the core Zn-binding fold. The upstream C4520 residue resides within this unstructured region, which together with flanking residues, forms a mobile region we term the mediator loop. The Zn coordination configuration (C5HC7HC2) of the TC domain is semi-contiguous and does not adopt cross-brace architecture (**Extended Data Figure 7c**).

Crystal packing revealed that T4380, within the N-terminus of a symmetry-related MYCBP2$_{cat}$ molecule (T4380$_{sym}$), was placed proximal to the esterification site where it forms a number of substrate-like interactions (**Figure 5a & b**). Firstly, the β-hydroxy group of T4380$_{sym}$ complements E4534 and H4583 and forms a potential triad (**Figure 5a)**. Thus the β-oxygen atom of T4380$_{sym}$ appears to be primed for deprotonation and nucleophilic attack. A catalytically productive electrophilic center is the C-terminus of Ub when thioester-linked to C4572. Even though this Ub molecule is absent in our structure, the C4572 sulfur atom is 3.8 Å away from the β-oxygen atom of T4380$_{sym}$. Thus, the structure appears to accurately reflect a catalytic intermediate poised to undergo threonine ubiquitination by esterification of its β-hydroxy group (**Figure 5a**). Furthermore, a sub-cluster of Phe residues (F4573, F4578 and F4586), proximal to the β-methyl group of T4380$_{sym}$, (**Figure 5a and b**) forms a well-defined hydrophobic pocket which the T4380$_{sym}$ β-methyl group docks into and seems to be a positive selectivity determinant for the threonine side chain. The proposed roles of these residues were validated in threonine discharge assays (**Figure 5c**). An H4583N mutation abolished activity consistent with a role as a general base. The H4583N mutant also underwent enhanced, thiol-sensitive, Ub adduct formation in accordance with the anticipated defect in rendering substrate nucleophiles reactive towards the C4572 thioester

(**Extended Data Figure 8a**). Conservative perturbation of the phenylalanine cluster also markedly reduced threonine discharge activity (**Figure 5c**). Perturbation of E4534 did not reduce activity hence its precise role remains unclear.

Conservation of RING domain binding to E2[16-18,23] permitted the modelling of an E2-RCR E3 ligase complex which was geometrically compatible with transthiolation between E2~Ub and MYCBP2$_{cat}$ C4520 **(Figure 5d and Extended Data Figure 8b**). To simulate the conformation required for subsequent Ub relay to C4572 we modelled the missing mediator loop residues with a GlyGly dipeptide thioester linked to C4520, representative of the Ub C-terminus that would be transferred via transthiolation with E2~Ub if our relay model was valid (**Extended Data Figure 8c-e**). In support of this mechanism the carbonyl C atom of the Ub thioester could be positioned in proximity (~3 Å) of the C4572 sulfhydryl sulfur atom. To adopt this conformation minor twisting of a GlyGly motif (residues 4515-4516) at the tip of the βA2 strand was necessary. Clashes were observed between mediator loop residues further C-terminal with R4533, E4534, N4580, H4583 and D4584 but these could largely be relieved by rotations of their sidechains into available space. As ordered loop residues 4527-4531 required a significant displacement to generate the model, we speculate that the mobile mediator loop region would span residues 4515-4531. As C4520, which resides within this mobile structural element, needs to be engaged by the E2 active site[24] this might account for the uncharacterized E2 residue requirements. An explanation for the inability to render the S4520 mutant catalytic in earlier experiments is its dynamic nature and the absence of a general base which could suppress the pK_a of the otherwise fully protonated S4520 side chain. Hence native C4520 catalytic activity is likely to arise from the sulfhydryl groups intrinsic nucleophilicity (**Extended Data Figure 8a**).

Although non-lysine ubiquitination has been reported[25-27], a human E3 ligase that preferentially carries out this function has remained elusive. Our characterization of the novel RCR E3 ligase found in MYCBP2 suggests that ubiquitination by esterification is intrinsic to higher eukaryotes and may be a regulator of synapse development and axon degeneration. Furthermore, non-protein Ub substrates (e.g. lipids, carbohydrates) have not been reported but considering the high esterification activity of MYCBP2 towards small molecule hydroxy compounds, this remains a possibility. It is not immediately clear why the proposed relay (**Figure 6a**) mechanism would have evolved. However, transthiolation is a cofactor independent process providing a facile means to shuttle Ub throughout the ubiquitin system[1]. We speculate that on steric grounds direct E2-E3 transthiolation with the structurally rigid and highly conserved E2 ubiquitin conjugating domain (Ubc)[28], and serine/threonine activity, are mutually exclusive at the esterification site and evolution of the mediator loop addresses this compatibility issue. Bioinformatic analysis revealed that MYCBP2 orthologues are found in virtually all animals but human homologues are unlikely to exist (**Extended Data Table 2**). Stabilization of NMNAT2 through MYCBP2 inhibition is a promising therapeutic strategy for mitigating neuron damage after injury and administration of chemotherapeutics[10,29,30] and slowing the progression of a range of neurodegenerative diseases including Alzheimer's and Parkinson's[29]. The delineation of this apparent Ub relay mechanism and the structural characterization of the molecular machinery responsible opens up new medical potential for treating a range of neurological conditions.

References

1 Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67**, 425-479, (1998).

2 Deshaies, R. J. & Joazeiro, C. A. RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78**, 399-434, (2009).

3 Huibregtse, J. M., Scheffner, M., Beaudenon, S. & Howley, P. M. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proc Natl Acad Sci U S A* **92**, 2563-2567 (1995).

4 Scheffner, M., Nuber, U. & Huibregtse, J. M. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature* **373**, 81-83, (1995).

5 Wenzel, D. M., Lissounov, A., Brzovic, P. S. & Klevit, R. E. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* **474**, 105-108, (2011).

6 Hewings, D. S., Flygare, J. A., Bogyo, M. & Wertz, I. E. Activity-based probes for the ubiquitin conjugation-deconjugation machinery: new chemistries, new tools, and new insights. *FEBS J* **284**, 1555-1576, (2017).

7 Pao, K. C. *et al.* Probes of ubiquitin E3 ligases enable systematic dissection of parkin activation. *Nat Chem Biol* **12**, 324-331, (2016).

8 Niphakis, M. J. & Cravatt, B. F. Enzyme inhibitor discovery by activity-based protein profiling. *Annu Rev Biochem* **83**, 341-377, (2014).

9 Borodovsky, A. *et al.* Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family. *Chem Biol* **9**, 1149-1159 (2002).

10 Zhen, M., Huang, X., Bamber, B. & Jin, Y. Regulation of presynaptic terminal organization by C. elegans RPM-1, a putative guanine nucleotide exchanger with a RING-H2 finger domain. *Neuron* **26**, 331-343 (2000).

11 Wan, H. I. *et al.* Highwire regulates synaptic growth in Drosophila. *Neuron* **26**, 313-329 (2000).

12 Grill, B., Murphey, R. K. & Borgen, M. A. The PHR proteins: intracellular signaling hubs in neuronal development and axon degeneration. *Neural Dev* **11**, 8 (2016).

13 Byrne, R., Mund, T. & Licchesi, J. Activity-Based Probes for HECT E3 ubiquitin ligases. *Chembiochem*, **18**, 1415-1427 (2017).

14 Stieglitz, B., Morris-Davies, A. C., Koliopoulos, M. G., Christodoulou, E. & Rittinger, K. LUBAC synthesizes linear ubiquitin chains via a thioester intermediate. *EMBO Rep* **13**, 840-846, (2012).

15 You, J. & Pickart, C. M. A HECT domain E3 enzyme assembles novel polyubiquitin chains. *J Biol Chem* **276**, 19871-19878 (2001).

16 Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115-120, (2012).

17 Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nat Struct Mol Biol* **19**, 876-883, (2012).

18 Lechtenberg, B. C. *et al.* Structure of a HOIP/E2~ubiquitin complex reveals RBR E3 ligase mechanism and regulation. *Nature* **529**, 546-550, (2016).

19 Dove, K. K. *et al.* Structural Studies of HHARI/UbcH7~Ub Reveal Unique E2~Ub Conformational Restriction by RBR RING1. *Structure* **25**, 890-900 e895, (2017).
20 Xiong, X. *et al.* The Highwire ubiquitin ligase promotes axonal degeneration by tuning levels of Nmnat protein. *PLoS Biol* **10**, e1001440, (2012).
21 Babetto, E., Beirowski, B., Russler, E. V., Milbrandt, J. & DiAntonio, A. The Phr1 ubiquitin ligase promotes injury-induced axon self-destruction. *Cell Rep* **3**, 1422-1429, (2013).
22 Milde, S., Gilley, J. & Coleman, M. P. Subcellular localization determines the stability and axon protective capacity of axon survival factor Nmnat2. *PLoS Biol* **11**, e1001539, (2013).
23 Zheng, N., Wang, P., Jeffrey, P. D. & Pavletich, N. P. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* **102**, 533-539 (2000).
24 Olsen, S. K. & Lima, C. D. Structure of a ubiquitin E1-E2 complex: insights to E1-E2 thioester transfer. *Mol Cell* **49**, 884-896, (2013).
25 Cadwell, K. & Coscoy, L. Ubiquitination on nonlysine residues by a viral E3 ubiquitin ligase. *Science* **309**, 127-130, (2005).
26 Shimizu, Y., Okuda-Shimizu, Y. & Hendershot, L. M. Ubiquitylation of an ERAD substrate occurs on multiple types of amino acids. *Mol Cell* **40**, 917-926, (2010).
27 Wang, X., Herr, R. A. & Hansen, T. H. Ubiquitination of substrates by esterification. *Traffic* **13**, 19-24, (2012).
28 Alpi, A. F., Chaugule, V. & Walden, H. Mechanism and disease association of E2-conjugating enzymes: lessons from UBE2T and UBE2L3. *Biochem J* **473**, 3401-3419, (2016).
29 Conforti, L., Gilley, J. & Coleman, M. P. Wallerian degeneration: an emerging axon death pathway linking injury and disease. *Nat Rev Neurosci* **15**, 394-409, (2014).
30 Ali, Y. O., Bradley, G. & Lu, H. C. Screening with an NMNAT2-MSD platform identifies small molecules that modulate NMNAT2 levels in cortical neurons. *Sci Rep* **7**, 43846, (2017).

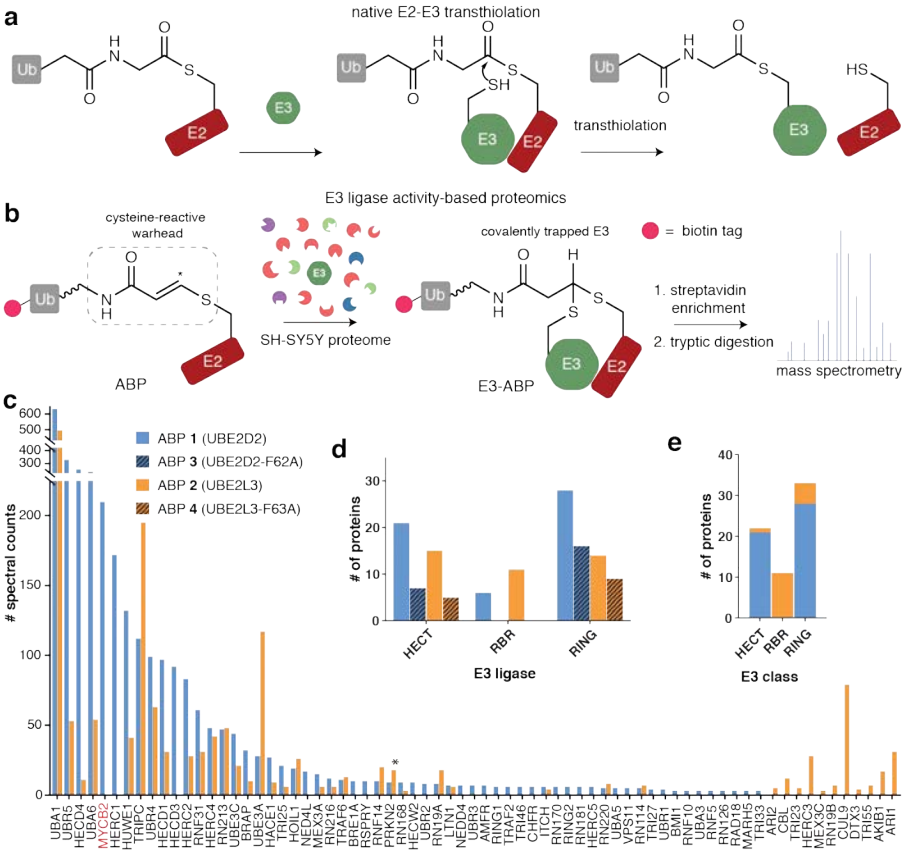


Figure 1. Activity-based Proteomics of E3 ligases. (**a**) Native transthiolation between E2~Ub and E3. (**b**) The ABP acts like a suicide substrate and covalently traps E3s that demonstrate transthiolation activity. Biotin-tagged ABPs enable parallelized MS-based HECT/RBR-like E3 profiling. Wavy bond represents triazole linkage to a truncated Ub molecule (see ref. 7, Extended Data Figure 1 and Methods). (**c**) Spectral counts obtained from ABP profiled OA-treated SH-SY5Y neuroblastoma cells for proteins currently annotated as E3s and E1s (**d**) Total number of HECT, RBR and RING E3s detected for the different probes. As the mutations introduced into control ABPs **3** and **4** are likely to impair rather than abolish E3 labelling, proteins with less than half the number of spectral counts of those obtained with their parental ABP (**1** and **2**, respectively) were not plotted. The asterisk highlights the RBR E3 ligase Parkin (PRKN2) which is not detected when oligomycin and antimycin A (OA)

treatment is withheld (**Extended Data Figure 2c, d**). (**e**) Number of E3 ligases detected using ABPs **1** and **2**.

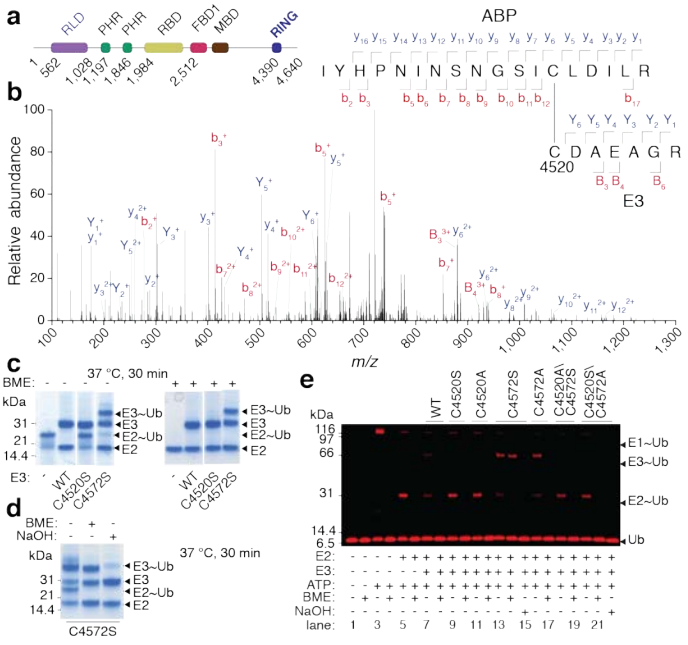


Figure 2. MYCBP2 is a novel class of E3 ligase and data support a cysteine relay mechanism. (**a**) Domain architecture of MYCBP2. Abbreviations: RCC1-like GEF domain (RLD), two PHR family-specific (PHR) domains, a RAE1 binding domain (RBD), a F-box binding domain 1 (FBD1), a Myc binding domain (MBD), and a C-terminal RING domain[10]. (**b**) Representative MS^2 spectrum for a crosslinked peptide derived from ABP treated WT $MYCBP2_{cat}$ illustrating that C4520 becomes labelled. Spectrum is for a 5^+ precursor ion (observed *m/z* = 614.5088; expected *m/z* = 614.5094). (**c**) Multiple turnover E2 discharge assay onto hydroxy nucleophiles (Tris and glycerol) present in the reaction buffer. **(d)** The Ub adduct on MYCBP2 C4572S is base labile (0.14 N NaOH) indicative of the formation of an engineered (oxy)ester linkage between Ub and the S4572 residue. **(e)** SDS-PAGE

thioester/ester trapping assay tracking Cy3b-labelled Ub. All biochemical data were consistent across three biological replicates.

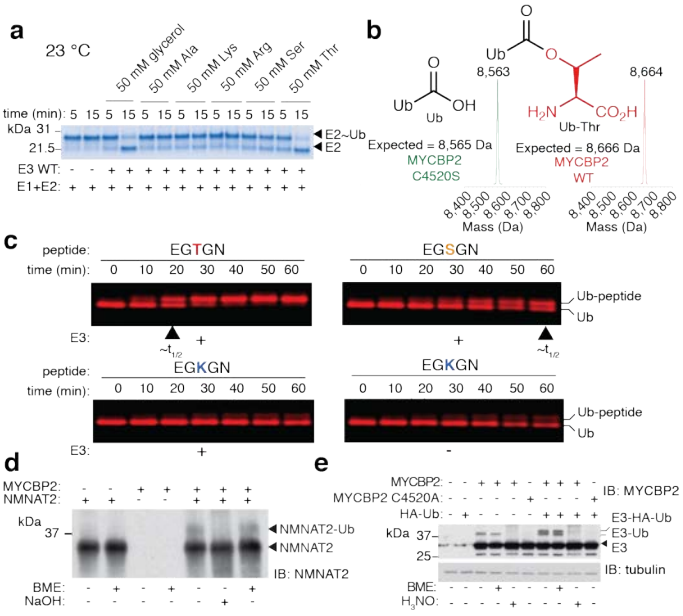


Figure 3. MYCBP2 ubiquitinates serine and threonine with selectivity for threonine (**a**) E3-mediated multiple turnover discharge reaction for a panel of amino acids (50 mM). Similar results obtained over 3 independent experiments. (**b**) Deconvoluted mass spectrum of Ub species after reaction with threonine (observed molecular weight of ubiquitinated threonine (Ub-Thr) = 8664 Da; theoretical = 8666 Da, observed molecular weight of unmodified Ub = 8563; theoretical = 8565 Da). N.B. Ub-Thr can undergo O-N acyl transfer forming the peptide-linked species. (**c**) SDS-PAGE analysis of peptide modification with fluorescently labeled Ub. Based on half-life approximation, MYCBP2$_{cat}$ is 3-fold selective for threonine over serine in the depicted peptide context. Similar results obtained over 2 independent experiments. (**d**) Recombinant NMNAT2 undergoes base-labile ubiquitination in a MYCBP2$_{cat}$ dependent manner. Similar results obtained over 2 independent experiments. (**e**) MYCBP2$_{cat}$ transiently transfected into HEK293 cells undergoes hydroxy

ubiquitination that is dependent on C4520. H_3NO corresponds to treatment with 0.5 M hydroxylamine at pH 9.0. Similar results obtained over 3 independent experiments.

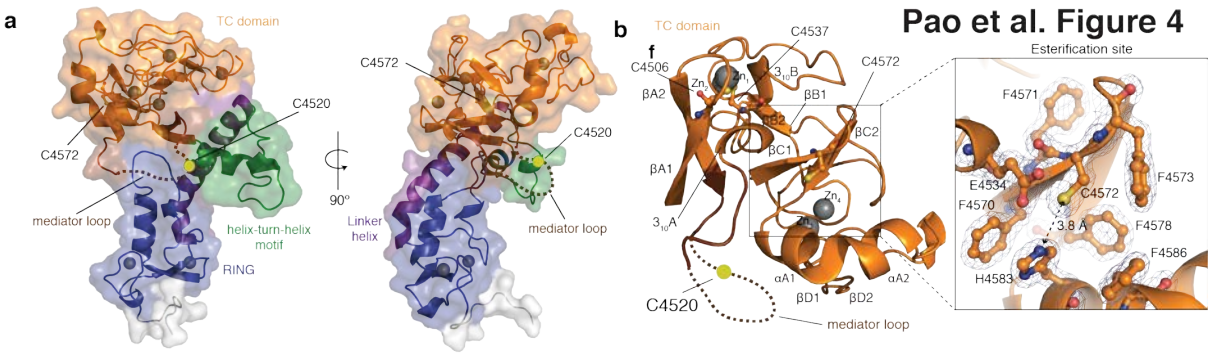


Figure 4. Crystal structure of MYCBP2$_{cat}$. (**a**) Surface and cartoon representation of the structure refined to 1.75 Å. (**b**) The tandem cysteine (TC) domain and its secondary structure annotation. Inset is a close up of the esterification site where a triad-like arrangement centers on C4572. A $2|F_{obs}|-|F_{calc}|$ electron density map contoured at 1.5 σ is represented in mesh for esterification site residues. C4506 and C4537, that abolished ABP labeling when mutated, are Zn-coordinating residues.

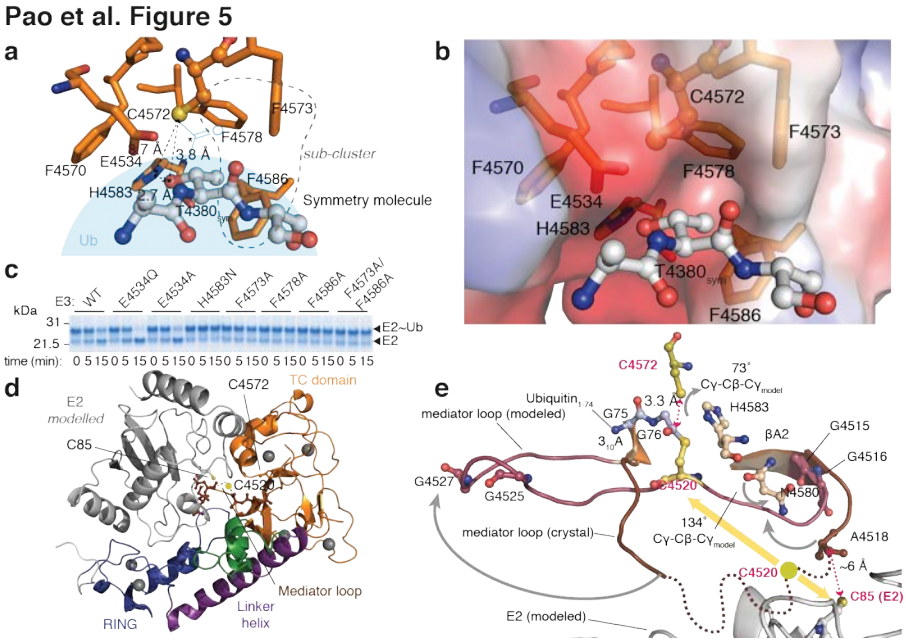


Figure 5. Structural basis for threonine selectivity, model of an E2-E3 intermediate and model of Ub relay. (**a**) A threonine residue at the N-terminus of a symmetry related molecule ($T4380_{sym}$) in an Ala-Thr-Ser sequence motif (grey ball and stick representation) is docked into the esterification site (orange stick). In light blue is a schematic of Ub thioester-linked to C4572. Asterisk corresponds to the electrophilic center of the Ub thioester carbonyl carbon atom. (**b**) Electrostatic potential of the esterification site (blue represents positive potential, red represents negative potential, and grey represents neutral potential). (**c**) Residues deemed important for catalysis were mutated and tested in threonine discharge assays. (**d**) Superposition of the RING domain from the RBR E3 ligase HOIP in complex with E2 (Ub linked to E2 has been omitted due to direct clash with the TC domain[18]; PDB ID: 5EDV) allows modelling of the E2 into our structure (grey cartoon representation). **(e)** Modeling of the proposed Ub relay mechanism. For the experimental structure, TC domain residues are in orange and mediator loop residues are in dark brown. For the model, TC domain residues are in light orange and mediator loop residues are in mauve. The modeled E2 is in grey cartoon. Ub residues (G75-G76) are in blue ball and stick.

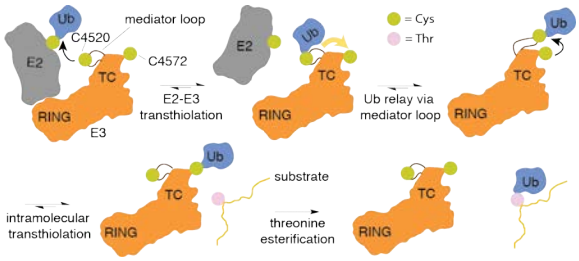


Figure 6. Schematic representation for the proposed model of the RCR E3 ligase mechanism. MYCBP2 has esterification activity towards serine and threonine but as it demonstrates preference for threonine, only this amino acid been depicted.

Acknowledgments

We thank D. Campbell and J. Varghese of the MRC PPU Proteomics Facility, MRC PPU DNA Sequencing Facility, the European Synchrotron Radiation Facility (ESRF), J. Hastie, H. McLauchlan and F. Brown from MRC PPU Reagents and Services for production of MYCBP2 antibody production, F. Zuccotto and I. Gilbert for software access, H. Walden and R. T. Hay for critical reading of the manuscript and valuable suggestions. This work was funded by the Scottish Funding Council, the UK Medical Research Council (MC_UU_12016/8), BBSRC (BB/P003982/1), and pharmaceutical companies supporting the Division of Signal Transduction Therapy (Boehringer-Ingelheim, GlaxoSmithKline and Merck KGaA). D.v.A. is funded by a Wellcome Trust Investigator Award (110061).

Author Contributions

S.V. and K.-C. P. designed research. K.-C. P. carried out experiments with assistance from S.V. N.T.W. performed cloning. A.K. prepared E1, E2 panel and recombinant NMNAT2 protein. K.R. mounted crystals and collected synchrotron diffraction data. M.S. contributed to project conception and carried out molecular synthesis. P.D.M. performed structural modeling, R.S. carried out SEC-MALS. K.H. performed bioinformatics analysis. D.v-A. solved the $MYCBP2_{cat}$ crystal structure. S.V. coded PERL scripts, processed MS data and wrote the manuscript with input from all authors.

Author Information

Reprints and permissions information is available at www.nature.com/reprints. Competing interests statement: S.V., K.-C.P. and M.S. are authors on patents relating to work presented in this article. Correspondence and requests for materials should be addressed to S.V. (s.s.virdee@dundee.ac.uk).

Materials and Methods

General materials

All DNA constructs were verified by DNA sequencing, (Medical Research Council Protein Phosphorylation and Ubiquitylation Unit, University of Dundee). DNA for bacterial protein expression was transformed into *E. coli* BL21-DE3 (Merck). All cDNA plasmids and antibodies generated for this study are available to request through our reagents website (https://mrcppureagents.dundee.ac.uk/). All solvents and reagents were purchased from Sigma-Aldrich or VWR unless otherwise stated.

Biotin functionalized ABP preparation

Ub with a GCSSG N-terminal extension was expressed from plasmid *pTXB1-UbΔ74-76-T3C* plasmid[7]. An equivalent plasmid encoding Ub residues 1-74 (*pTXB1-UbΔ75-76-T3C*) was also created. Ub thioesters were obtained as described previously generating cysteine tagged Cys-Ub_{1-73}-SR and Cys-Ub_{1-74}-SR, respectively[7]. The extended Ub_{1-74} was included as this retains Arg74 which forms a favorable electrostatic interaction with the RBR E3 HOIP[31]. Cys-Ub_{1-73}-SR (30 mg) was reconstituted by the addition of DMSO (116 µL) followed by H_2O (456 µL). An aqueous stock solution (48 mM) of EZ-link Iodoacetyl-PEG2-biotin (Thermofisher) was prepared and 200 µL was added to the Cys-Ub_{1-73}-SR solution (580 µl) followed by the addition of 900 µl degased buffer (50 mM Na_2HPO_4 pH 7.5, 150 mM NaCl). The reaction was incubated at 23 °C for 1 hour and monitored by LC-MS. The protein (Biotin-Ub_{1-73}-SR) was then further purified by semi-preparative RP-HPLC (Column: BioBasic-

4; Part number: 72305-259270). A gradient of 20 % buffer A to 50 % buffer B was applied at a flow rate of 10 mL min^{-1} over 60 min (buffer A=0.1 % TFA in H_2O, buffer B=0.1 % TFA in acetonitrile). The above procedure was repeated to generate Biotin-Ub_{1-74}-SR. HPLC fractions containing Biotin-Ub_{1-7X}-SR were pooled and lyophilized (Yield: Biotin-Ub_{1-73}-SR 75-85 %, Biotin-Ub_{1-74}-SR 40-50 %) (**Extended Data Figure 1a and d**). Biotin-tagged ABPs containing thioacrylamide warheads were then prepared as previously described[7] employing the E2 recognition elements UBE2D2*, UBE2D2* F62A, UBE2L3* and UBE2L3* F63A furnishing ABPs **1**, **3**, **2** and **4** (**Extended Data Figure 1b, c, e and f**). *denotes E2 where non-catalytic Cys residues were mutated to Ser. ABPs based on UBE2D2* and UBE2D3* bearing hexahistidine reporter tags and thioacrylamide warheads (**Extended Data Figure 3a**), were also prepared yielding ABPs **5** and **6**, respectively.

Cell culture and lysis Protocol

SH-SY5Y cells were cultured as previously described[7]. HEK293 were cultured (37 °C, 5 % CO_2) in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % (v/v) Fetal Bovine Serum (FBS), 2.0 mM L-glutamine, and antibiotics (100 units mL^{-1} penicillin, 0.1 mg mL^{-1} streptomycin). Cell transfections were performed using polyethylenimine (Polysciences) according to the manufacturer's instruction. MG-132 (50 µM) was added to cells two hours prior to lysis. Cells were rinsed with ice-cold PBS and extracted in lysis buffer (1 % NP-40, 50 mM Tris-HCl pH 7.5, 1 mM EGTA, 1 mM EDTA, 0.27 M sucrose, 10 mM sodium 2-glycerophosphate, 0.2 mM phenylmethane sulfonyl fluoride (PMSF), 1.0 mM benzamidine, 1 mM sodium ortho-vanadate, 50 mM sodium fluoride and 5 mM sodium pyrophosphate, 50 mM iodoacetamide and cOmplete™, EDTA-free protease inhibitor

cocktail (Roche)). Lysates were then clarified by centrifugation at 4 °C for 30 min at 21,100 x g rpm. Supernatants were collected (total cell extracts) and protein concentration determined by Bradford assay. For the base-lability test, indicated cell lysates were further incubated with 0.5 M hydroxylamine, pH 9.0 at 37 °C for 30 minutes. Mycoplasma tests were carried out in accordance with departmental protocols and tested negative.

Immunoblotting

Samples were mixed with NuPAGE LDS sample buffer (Thermofisher) without boiling, and resolved by SDS–PAGE (4–12 % NuPage gel, Thermofisher) with MOPS or MES running buffer and transferred on to 0.45 µm nitrocellulose membranes (GE Life Sciences). Membranes were blocked with PBS-T buffer (PBS + 0.1 % Tween-20) containing 5 % (w/v) non-fat dried skimmed milk powder (PBS-TM) at room temperature for 1 h. Membranes were subsequently probed with the indicated antibodies in PBS-T containing 5 % (w/v) Bovine Serum Albumin (BSA) overnight at 4 °C. Detection was performed using HRP-conjugated secondary antibodies in PBS-TM for 1 h at 23 °C. ECL western blotting detecting reagent (GE Life Sciences) was used for visualization in accordance with the manufacturers protocol.

Antibodies

His-tagged species were probed with 1:10000 anti-His primary antibody (Clontech, #631212). Alpha tubulin (1E4C11) mouse mAb (Proteintech®) was used at 1:10000 dilution. The MYCBP2 antibody was used at 0.5 µg mL^{-1} and raised in sheep by MRC PPU Reagents and Services and affinity-purified against the indicated antigen: anti-MYCBP2 (2nd bleed of SA357,

residues 4378-4640 of human MYCBP2). Mouse monoclonal NMNAT2 antibody (clone 2E4; Sigma Aldrich) was used at 0.5 μg mL^{-1}.

Activity-based proteomic profiling of SH-SY5Y cells

SH-SY5Y total cell lysate (4.5 mg, 550 μL) was mixed with ABPs **1**, **2**, **3** and **4** (3 μM) and incubated at 30 °C for 4 hours. To induce Parkin activation cells were administered with oligomycin (5 μM) and antimycin A (10 μM) (OA) for 3 hours. Control enrichments were also performed where probe was withheld. Extracts were mixed with 100 μL of Pierce™ Streptavidin Plus UltraLink™ Resin (ThermoFisher Scientific) and diluted with 6 % SDS solution (20 μL) to a final concentration of 0.2 %. Samples were incubated for 4 hours at 4 °C and resin washed (2 mL 0.2 % SDS/PBS, 2 mL PBS, 1 mL 4 M urea/PBS, 2 ml PBS) and then resuspended in 190 μL Tris buffer (50 mM Tris pH 8, 1.5 M urea). Resin-bound proteins were reduced with TCEP (5 mM) for 30 minutes at 37 °C and then alkylated with iodoacetamide (10 mM) at 23 °C for 20 minutes. DTT (10 mM) was then added followed by washing with buffer (50 mM Tris pH 8, 1.5 M urea) to a final volume of 300 μL. Trypsin (2 μg) was then added and further incubated at 37 °C for 14 hours. Trifluoroacetic acid was added to a final concentration of 0.1 % and samples were desalted with a C^{18} MacroSpin column (The Nest Group Inc). LC-MS/MS analysis was performed on an LTQ Orbitrap Velos instrument (Thermo Scientific) coupled to an Ultimate Nanoflow HPLC system (Dionex). A gradient running from 3 % solvent B to 99 % solvent B over 345 min was applied (solvent A = 0.1 % formic acid and 3 % DMSO in H_2O; solvent B = 0.08 % formic acid and 3 % DMSO in 80 % MeCN).

Data processing

Raw files were searched against the Swissprot database and a decoy database using the MASCOT server (Matrix Science). Trypsin specificity with up to three missed cleavages was applied. Cysteine carbamylation was set as a fixed modification and variable modifications were methione oxidation/dioxidation. A PERL script was used to extract the number of rank 1 peptides for each protein from the MASCOT search results and this figure was used as the number of spectral counts. A second PERL script filtered the data by searching the human swisspfam_v30 database using the E3 domain terms RING, HECT, IBR and zf-UBR. Manual curation was also carried out which involved the addition of E1 enzymes. Any proteins with less than 3 spectral counts and less than 14-fold spectral count enrichment, relative to control experiments where ABP was withheld, were omitted from the list. Pairwise datasets were then plotted as column charts in Prism (Graphpad Software).

Cloning of MYCBP2$_{cat}$

Human MYCBP2 (NM_015057.4) sequences were amplified from full-length Addgene plasmid #2570. Wild type and mutant fragments were subcloned as BamHI/Not1 inserts into pGEX6P-1 (GE Life Sciences) for bacterial expression, or a modified version of pcDNA TM5/FRT/TO (ThermoFisher) containing an N-terminal Myc tag for mammalian expression.

UBE1 and E2 expression and purification

6His-UBE1 was expressed in Sf21 cells and purified via its tag as previously described[32]. Phosphate Buffered Saline was used throughout the purification and hydroxy containing

compounds avoided. UBE2D3 was expressed as an N-terminally 6His-tagged protein in BL21 cells and purified over Ni-NTA-agarose and dialysed into 50 mM Na_2HPO_4 pH 7.5, 150 mM NaCl, 0.5 mM TCEP. UBE2A was expressed as a GST-fusion in *E. coli* and the GST tag was proteolytically removed. The remaining E2s were expressed as recombinant bacterial proteins and purified via their His-tags and buffer exchanged by size exclusion chromatography into running buffer (50 mM Na_2HPO_4 pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.015 % Brij-35) using a Superdex 75 column (GE Life Sciences).

Expression and purification of MYCBP2 and GST-MYCBP2

GST-tagged $MYCBP2_{cat}$ (Ser4378-Phe4640), WT and mutants, were expressed at 16 °C overnight and purified against glutathione resin (Expedeon) using standard procedures. GST-tagged constructs were eluted with glutathione and untagged constructs were obtained by on-resin cleavage with Rhinovirus 3C protease. Proteins were buffer exchanged into storage buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1.0 mM TCEP) and kept at -80 °C.

Expression and Purification of NMNAT2

NMNAT2 was expressed with a 6His-SUMO tag in BL21(DE3) cells, induced with 0.1 mM IPTG and incubated for expression at 16 °C. The cells were collected and lysed in buffer (50 mM Tris-HCl (pH 7.5), 250 mM NaCl, 0.2 mM EGTA, 20 mM imidazole, 20 mM L-arginine, 0.015 % Brij 35, 1 mM Leupeptin, 1 mM Pefabloc, 1 mM DTT using standard protocols and the protein was purified over Ni-NTA-agarose. The eluted protein was incubated with His-SENP1 protease during dialysis against PBS, 20 mM L-arginine, 1 mM DTT. The tag and protease were

depleted against Ni-NTA-agarose and NMNAT2 was concentrated and subjected to chromatography on a Superdex 75 HR 10/30 into buffer (PBS, 20 mM L-arginine).

Activity-based protein profiling of MYCBP2 cysteine mutants

The indicated MYCBP2 mutant was diluted into Tris buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl) to a final concentration of 3 µM. Probe **6** was added (12 µM) and incubated with E3 ligase at 30 °C for 4 hours. Reactions were quenched by the addition of 4X LDS loading buffer (supplemented with ~680 mM 2-mercaptoethanol) and samples were resolved by SDS-PAGE (4-12 % NuPage gel) followed by Coomassie staining or anti-His immunoblotting.

Tryptic MS/MS sequencing of probe-labelled MYCBP2

Crosslinking MS using ABP **6** was carried out as previously described[7]. In summary the Coomassie stained SDS-PAGE band corresponding to ABP-labelled WT MYCBP2 was analyzed by LC-MS/MS using an Orbitrap Fusion™ Tribrid™ mass spectrometer (Thermo Scientific) coupled to an Ultimate Nanoflow HPLC system (Dionex). A gradient running from 0 % solvent A to 60 % solvent B over 120 min was applied (solvent A = 0.1 % formic acid in H_2O; solvent B = 0.08 % formic acid in 80 % MeCN). Fragment ions were generated by HCD and 1^+, 2^+ and 3^+ precursor ions excluded. Raw data was searched using the pLink software[33] against UBE2D3* and MYCBP2 sequences with trypsin specificity (up to 2 missed cleavages). The error window for MS/MS fragment ion mass values was set to the software default of 20 ppm. A crosslinker monoisotopic mass of 306.1805 Da was manually added which accounted for the theoretical

mass difference associated with formation of a bisthioether between 2 Cys residues derived from probe **6,** which was based on UBE2D3* and contained a thioacrylamide AVS warhead[7].

Tris/glycerol-mediated E2 discharge assay

Assays were carried out in buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 5 mM $MgCl_2$) containing the indicated MYCBP2 mutant (15 µM), UBE1 (1.5 µM), UBE2D3 (15 µM), Ub (37 µM) and ATP (10 mM). The reactions were incubated at 37 °C for 30 minutes. Reactions were terminated by the addition of 4X LDS loading buffer (with and without ~680 mM 2-mercaptoethanol). A C4572S sample was further incubated with 0.14 N NaOH at 37 °C for 20 minutes and samples were resolved by SDS-PAGE (4-12 % NuPage gel) and visualized by Coomassie staining.

LC-MS analysis of nucleophile discharge assays

Reactions were prepared as described for discharge assay. After 30 minutes, the reaction was analyzed using an Agilent 1200/6130 LC-MS system (Agilent Technologies) using a 10-75 % gradient over 20 minutes (buffer A = H_2O + 0.05 % TFA, buffer B = acetonitrile + 0.04 % TFA).

Preparation of Cy3b-Ub

Ub bearing a TEV protease-cleavable N-terminal hexahistidine tag followed by a ACG motif was expressed in bacteria from a pET plasmid (kindly provide Ronald Hay, University of

Dundee). Protein was purified by Ni affinity chromatography, cleaved from the tag with TEV protease then buffer exchanged into reaction buffer (50 mM HEPES, pH 7.5, 0.5 mM TCEP). Protein was concentrated to 2 mg mL^{-1} and 221 μL (50 nmol) was mixed with Cy3b-maleimide (150 nmol, GE Life Sciences) in a final volume of 300 μL and agitated for 2 h at 25 °C. Labelled protein was then further purified with a P2 Centri-Pure desalting column (EMP Biotech) with degassed buffer (50 mM Na_2HPO_4, 150 mM NaCl).

MYCBP2 thioester/ester trapping assay

UBE1 (2 μM) was mixed with Cy3b-Ub (1 μM) in buffer (40 mM Na_2HPO_4-HCl pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 5 mM $MgCl_2$) (Figure 2e, *Lane 1, 2*). The reaction was then initiated by the addition of ATP (5 mM) and incubated for 10 minutes at 25°C. Samples (*lane 3*, *4)* were taken and combined with UBE2D3 (10 μM). After a further 10 minutes at 25°C, samples (lane 5, 6) were taken and combined with GST-MYCBP2$_{cat}$ (WT, C4520S, C4520A, C4572S, C4572A, C4520A/C4572S or C4520S/C4572A) (15 μM). The reactions were incubated at 25 °C for 30 seconds and terminated by the addition of 4X LDS loading buffer (either non-reducing or reducing). For Ub~GST-MYCBP2$_{cat}$ C4572S ester bond cleavage, 0.14 N NaOH was added after E3 reaction with E1/E2 mixture for 30 seconds and then further incubated at 37 °C for 20 minutes. The gel was then scanned with a Chemidoc Gel Imaging System (BioRad).

Multiple turnover amino acid and peptide panel discharge assays

Stock solutions (0.5 M) of amino acids were dissolved in MQ water and pH was adjusted to pH ~8. Peptides of the sequence Ac-EGXGN-NH_2 (X= K, S or T) were obtained from

Bio-Synthesis Inc. Stock peptide solutions (200 mM) were dissolved in MQ water and pH was adjusted to pH ~8. An E2 (UBE2D3) charging reaction was carried out in buffer (40 mM Na_2HPO_4-HCl pH 8.0, 150 mM NaCl, 0.5 mM TCEP) containing UBE1 (250-500 nM), UBE2D3 (20 μM), Ub (50 μM), or Cy3B-Ub (25 μM), $MgCl_2$ (5 mM) and ATP 10 (mM). The reaction was incubated at 37 °C for 15 minutes and then equilibrated to 23 °C for 3 minutes. An equivalent volume of nucleophile sample containing small molecule/peptide nucleophile (100 mM) and GST-MYCBP2 (10 μM) was then added and incubated at 23 °C. Samples were taken at the specified time points and analyzed as described for Tris/glycerol-mediated E2 discharge assay**.**

Cy3B-Ub was visualized using a Chemidoc Gel Imaging System (Biorad). LC-MS was carried out as described for Tris/glycerol discharge but amino acid substrate samples were quenched by the addition of 2:1 parts quenching solution (75 % acetonitrile, 2 % TFA) and peptide substrate samples were quenched by addition of 1:1 parts quenching solution.

Multiple turnover E2 discharge panel

E2s were screened for threonine discharge activity with GST-MYCBP2$_{cat}$ as described for the amino acid panel. E2s were also incubated in the presence of threonine but in the absence of GST-MYCBP2$_{cat.}$ These samples provided a reference to distinguish between intrinsic E2~Ub instability and E3-dependent discharge.

Single turnover E2 mutant discharge by in-gel fluorescence

E2 mutants[16,17,34-36] (10 μM) were charged with Cy3b-labelled Ub (12.5 μM) in a final volume of 12 μL at 37 °C for 20 minutes then cooled at 23 °C for 3 minutes. E2 recharging was

then blocked by the addition of MLN4924 derivative, Compound 1 (25 μM)[37], which inhibits E1, and then incubated for a further 15 minutes. The mixture was then mixed with 12 μL of GST-MYCBP2$_{cat}$ (5 μM) and threonine (100 mM) and incubated at 23 °C for the specified time. Analysis was carried out as for multiple turnover assays. To account for intrinsic E2~Ub instability the mean % discharge (n=2) calculated against a parallel incubation where E3 was withheld. Data were plotted using Prism (Graphpad).

Expression and purification of ARIH1 and UBE3C

ARIH1 residues 1-394 (Dundee clone DU24260) was expressed as an N-terminally GST-tagged fusion protein in BL21 cells. UBE3C residues 641 – 1083 (Dundee Clone DU45301) was expressed as a N-terminally GST-tagged fusion protein in Sf21 cells using a baculovirus infection system.

Calculation of observed rate constants for E3-substrate dependent single turnover E2~Ub discharge

UBE2D3 or UBE2L3 (5 μM) were charged with Cy3b-labelled Ub (8 μM) in a final volume of 30 μL at 37 °C for 25 minutes then incubated at 23 °C for 3 minutes. Single turnover conditions for E2~Ub discharge were achieved by E1 inhibition with MLN4924 derivative, Compound 1 (25 μM) and then incubated for a further 15 minutes. The mixture was then mixed with 30 μL of MYCBP2$_{cat}$ or ARIH1$_{1-394}$ (HHARI) or UBE3C$_{641-1083}$ (1 μM) and threonine (100 mM) and incubated at 23 °C for the specified time. Samples were quenched with non-reducing 4X LDS loading buffer and resolved by SDS-PAGE (Bis-Tris 4-12 %). The gel was then

scanned with a Chemidoc Gel Imaging System (BioRad) and subsequently Coomassie stained. E2~Ub signals were quantified using the Fiji software. Observed rate constants were obtained by fitting reaction progressive curves to a single exponential function using Prism (Graphpad Software).

MYCBP2 crystallization

MYCBP2 was expressed as described for untagged protein. After protease cleavage of the tag the protein was further purified by size exclusion chromatography using an ÄKTA FPLC system and a HiLoad 26/600 Superdex 75 pg column (GE Life Sciences). The running buffer consisted of 20 mM HEPES pH 7.4, 150 mM NaCl, 4 mM DTT. Combined fractions were concentrated to 10.4 mg mL^{-1}. Sparse matrix screening was carried out and Bipyrimidal crystals were obtained from the Morpheus screen condition C1 (Molecular Dimensions). A subsequent optimization screen yielded multiple crystals (Buffer system 1 (MES/imidazole) pH 6.7, 23.3 mM Na_2HPO_4, 23.3 mM $(NH_4)_2SO_4$, 23.3 mM $NaNO_3$, 18 % PEG500 MME, 9 % PEG20000). A single crystal was soaked in mother liquor and further cryoprotected by supplementation with 5 % PEG400 and frozen in liquid N_2. Data were collected to 1.75 Å at the European Synchrotron Radiation Facility at Beamline ID23-1. Energy was set to the peak value of 9.669 keV (1.2823 Å), as determined by an absorption edge energy scan. A total of 360 ° were collected with an oscillation range of Ω = 0.1 °. The phase problem was solved by locating 6 Zn^{2+} sites in the anomalous signal and solvent flattening with the SHELX suite. An initial model was built by ARP/wARP[38] and subsequently optimized by manual building in COOT[39] and refinement with REFMAC5[40] resulting in the final model with statistics as shown

in **Extended Data Figure 7**. Final Ramachandran statistics were favored: 95.55 %, allowed: 3.24 %, outliers: 1.21 %.

Size exclusion chromatography with multi-angle light scattering (SEC-MALS)

SEC–MALS experiments were performed on a Ultimate 3000 HPLC system (Dionex) with an in-line miniDAWN TREOS MALS detector and Optilab T-rEX refractive index detector (Wyatt). In addition, the elution profile of the protein was also monitored by UV absorbance at 280 nm. A Superdex 75 10/300 GL column (GE Life Sciences) was used. Buffer conditions were 50 mM Na_2HPO_4 pH 7.5, NaCl 150 mM, 1.0 mM TCEP and a flow rate of 0.3 mL min^{-1} was applied. Sample (50 μL, 5.5 mg mL^{-1}) was loaded onto the column with a Dionex autosampler. Molar masses spanning elution peaks were calculated using ASTRA software v6.0.0.108 (Wyatt).

Mediator loop modeling

Mediator loop residues where built and geometry optimized within the Bioluminate Software (Schrödinger). Side chains were modified within COOT[39] and figures were generated with Pymol (Schrödinger). Ramachandran analysis was carried with the RAMPAGE server[41].

NMNAT2 ubiquitination assay

NMNAT2 (5 µM) was mixed with E1 (500 nM), UBE2D3 (10 µM), $MYCBP2_{cat}$ (10 µM), Ub (50 µM), ATP (10 mM) and made up with 10X pH 7.5 buffer (40 mM $Na_2H_2PO_4$ pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM TCEP). The reactions were incubated at 37 °C for 1 hour and terminated by the addition of 4X LDS loading buffer (either non-reducing or reducing). For base lability test, reactions were supplemented with 0.14 N NaOH and then further incubated at 37 °C for 20 minutes.

Bioinformatic analysis

Proteins belonging to the RCR family were identified by generalized profile searches. Overall 671 such sequences were identified. The sequences were aligned by profile-guided alignment using the pftools package. For identifying representative sequences from various taxa, the Belvu program (Sanger Institute) was used to remove sequences with >80 % identity to other sequences. Truncated and misassembled proteins were removed manually, resulting in 130 representative TC domain sequences.

Code availability statement

Custom PERL scripts for processing of MS data are available upon request and should be addressed to S.V. (s.s.virdee@dundee.ac.uk).

Data availability statement

All data supporting the findings within this manuscript are within the main figures, extended data or supplementary files and information. Coordinates have been deposited with the Protein Data Bank (PDB ID 5O6C). For gel source data, see Supplementary Figure 1.

Methods References

31 Stieglitz, B. *et al.* Structural basis for ligase-specific conjugation of linear ubiquitin chains by HOIP. *Nature* **503**, 422-426, (2013).

32 Stanley, M. *et al.* Orthogonal thiol functionalization at a single atomic center for profiling transthiolation activity of E1 activating enzymes. *ACS Chem Biol* **10**, 1542-1554, (2015).

33 Yang, B. *et al.* Identification of cross-linked peptides from complex samples. *Nat Methods* **9**, 904-906, (2012).

34 Pruneda, J. N. *et al.* Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Mol Cell* **47**, 933-942, (2012).

35 Wu, P. Y. *et al.* A conserved catalytic residue in the ubiquitin-conjugating enzyme family. *EMBO J* **22**, 5241-5250, (2003).

36 Yunus, A. A. & Lima, C. D. Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. *Nat Struct Mol Biol* **13**, 491-499, (2006).

37 Brownell, J. E. *et al.* Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ. *Mol Cell* **37**, 102-111, (2010).

38 Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**, 1171-1179, (2008).

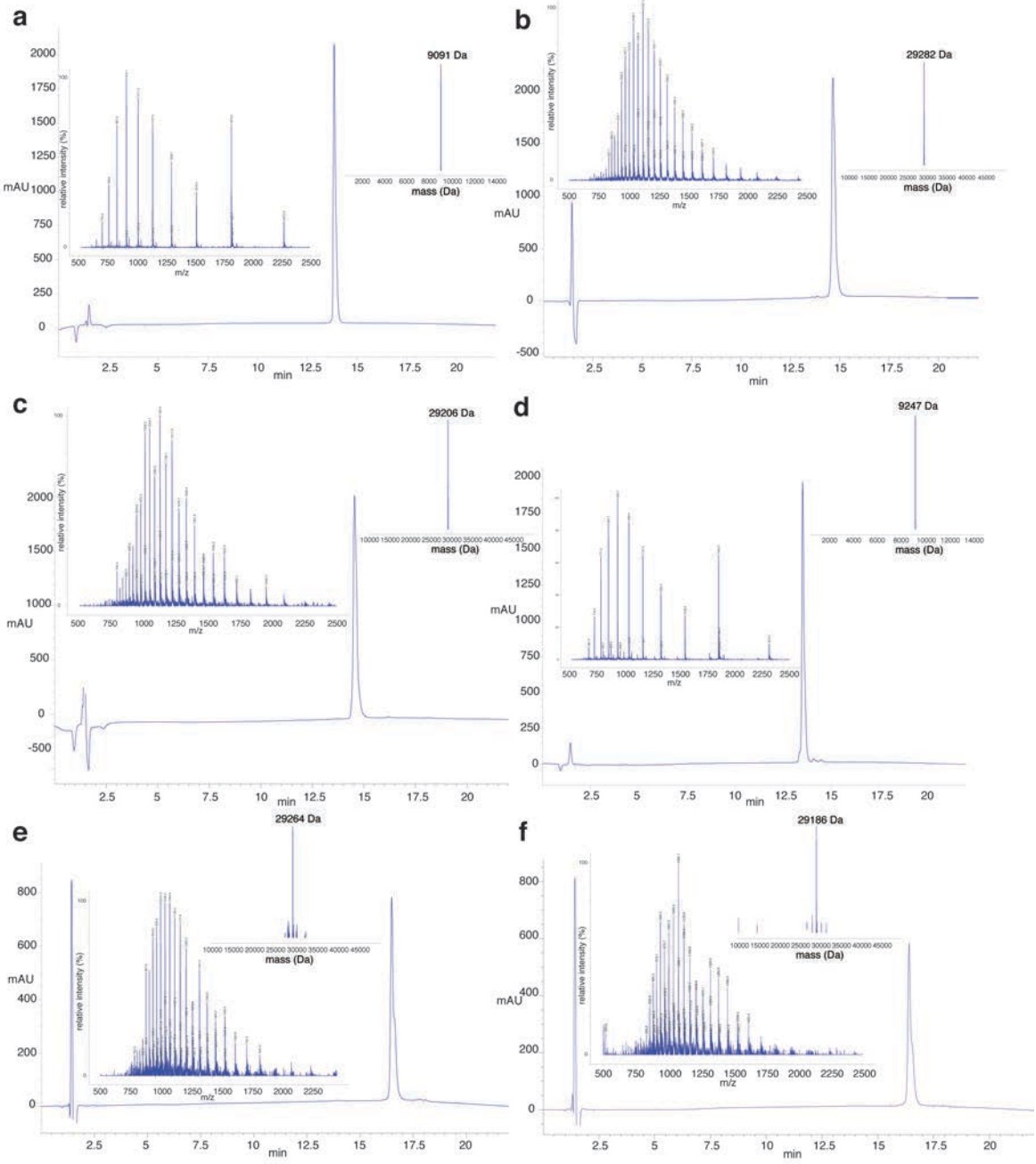
39 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, (2010).

40 Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367, (2011).

41 Lovell, S. C. *et al.* Structure validation by Cα geometry: phi, psi and Cβ deviation. *Proteins* **50**, 437-450, (2003).

Extended Data

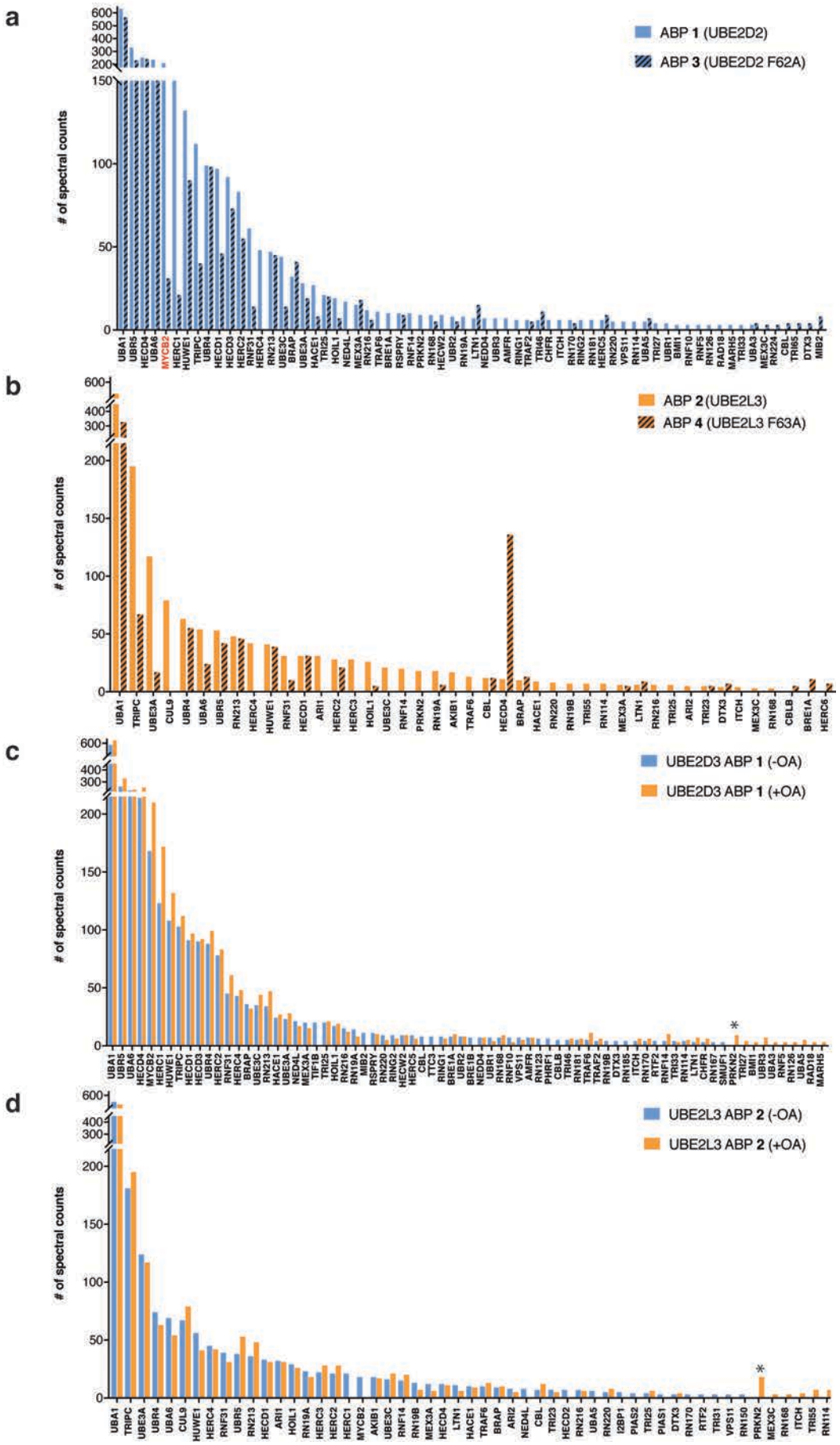
Extended Data Figure 1



Extended Data Figure 1. LC-MS characterization of biotinylated ABP intermediates and biotinylated ABPs. E3s can have distinct E2 preferences[5], so to obtain broad coverage we prepared biotinylated ABPs based on the promiscuous E2, UBE2D2, (UBE2D2-ABP **1**), and the HECT/RBR-specific E2, UBE2L3, (UBE2L3-ABP **2**). As controls we prepared ABPs

containing a point mutation in the E2 recognition component predicted to disrupt or impair E3 ligase binding (UBE2D2-F62A-ABP **3** and UBE2L3-F63A-ABP **4**, respectively)[7]. ABPs **1** and **3** were constructed with Ub extended by a single residue relative to that previously reported[7] (Ub_{1-74} rather than Ub_{1-73}) as this improved labelling efficiency of the RBR E3 HOIP. (a) Characterization of biotin labelled, truncated Ub thioester intermediate, Biotin-Ub_{1-73}-SR used for ABP **2** and **4**. HPLC chromatogram monitoring UV absorbance at 214 nm (as for all subsequent intermediates and ABPs). ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 9093 (-Met); found = 9091 Da. (b) Characterization of UBE2L3 ABP **2.** ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 29286.9 (-Met); found = 29282 Da. (c) Characterization of UBE2L3 F63A ABP **4.** ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 29210.8 (-Met); found = 29206 Da. (d) Characterization of probe intermediate with extended Ub C-terminus, Biotin-Ub_{1-74}-SR used to make ABP **1** and **3**. ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 9249.2 (-Met); found = 9247 Da. (e) Characterization of UBE2D2 ABP **1.** ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 29268.8 Da (-Met); found = 29264 Da. (f) Characterization of UBE2D2 F62A ABP **2.** ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 29192.7 Da (-Met); found = 29186 Da. Intermediates and probes have been prepared and characterized >3 times with similar results.

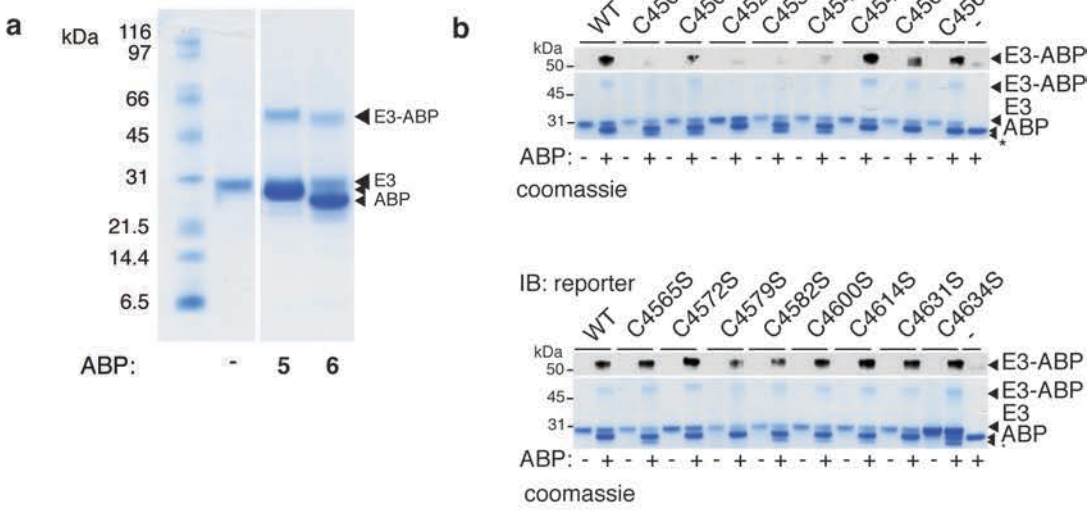
Extended Data Figure 2



Extended Data Figure 2. Activity-based proteomic profiling of neuroblastoma SH-SY5Y cells. Parallel profiling of neuroblastoma SH-SY5Y cell extracts was carried out with ABPs **1**-**4.** As an additional control, cells were left untreated or treated with inhibitors of oxidative phosphorylation, oligomycin and antimycin A (OA), which enables activity-dependent labelling of the RBR E3 Parkin[7]. ABP-labelled proteins were enriched against streptavidin resin followed by on-resin tryptic digestion. Obtained peptides were analyzed by data-dependent LC-MS/MS. Recovered proteins were filtered against E3-associated PFAM domain terms (RING, HECT, IBR, zf-UBR) and proteins with <3 spectral counts were excluded. E3s that did not demonstrate >14-fold spectral count enrichment compared to control purifications, where ABP was withheld, were also excluded. E1s yielded a strong signal as these undergo transthiolation and are highly enriched by our ABPs. The number of spectral counts for the majority of HECT/RBR E3s was reduced by > 50 % when the binding-defective control probes **3** and **4** were employed relative to their parental counterpart (**Figure 1c**). The aggregate number of recovered HECT/RBR E3s from probes **1** and **2** was 33 (22 HECT and 11 RBR) thus representing ~80 % of the currently annotated HECT/RBR E3s (**Figure 1c**). For the E3s that remain permissive to control ABPs **3** and **4** we cannot establish whether this is because the respective E3s are labelled in an activity-independent manner or whether they are permissive to the F62A/F63A mutation. Furthermore, ABP-dependent spectral count signals are not normalized against protein abundance. Hence, we cannot deconvolute the effects of E3 activation stoichiometry from E3 abundance (i.e. highly abundant E3s that are in a low activation state could yield disproportionately high signals in our data). **(a)** The number of spectral counts for the recovered E1/E3 proteins plotted against protein ID for UBE2D2 ABP **1** and the respective control probe UBE2D2 F62A ABP **3.** MYCBP2/Phr1 yields a high signal with ABP **1** and the signal is reduced >50% when the

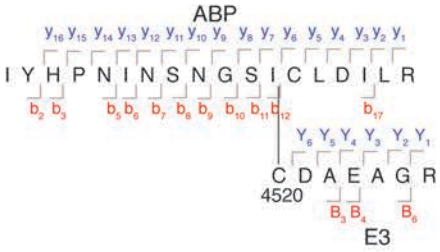
respective control ABP **3** is employed. A number of RING E3s, that will bind the ABP, are also labelled and for the majority of cases this is presumed to be mechanistically off-target labeling that is exacerbated by the high sensitivity of MS-based detection. Another possibility is that hitherto undiscovered RING-linked E3s are being labelled. **(b)** The number of spectral counts for the recovered E1/E3 proteins plotted against protein ID for UBE2L3 ABP **2** versus the respective control probe UBE2L3 F63A ABP **4**. MYCBP2 is not detected with UBE2L3 ABPs **2** and **4**. **(c)** The number of spectral counts for the majority of E1/E3 proteins obtained with ABP **1** for untreated versus OA-treated cells. **(d)** The number of spectral counts for the majority of E1/E3 proteins obtained with ABP **2** for untreated versus OA-treated cells. Parkin peptides were only recovered from cells treated with OA, consistent activity-dependent Parkin labelling. Taken together, this demonstrates that for at least a subset of detected E3s, spectral counts correlate with E3 activity.

Extended Data Figure 3



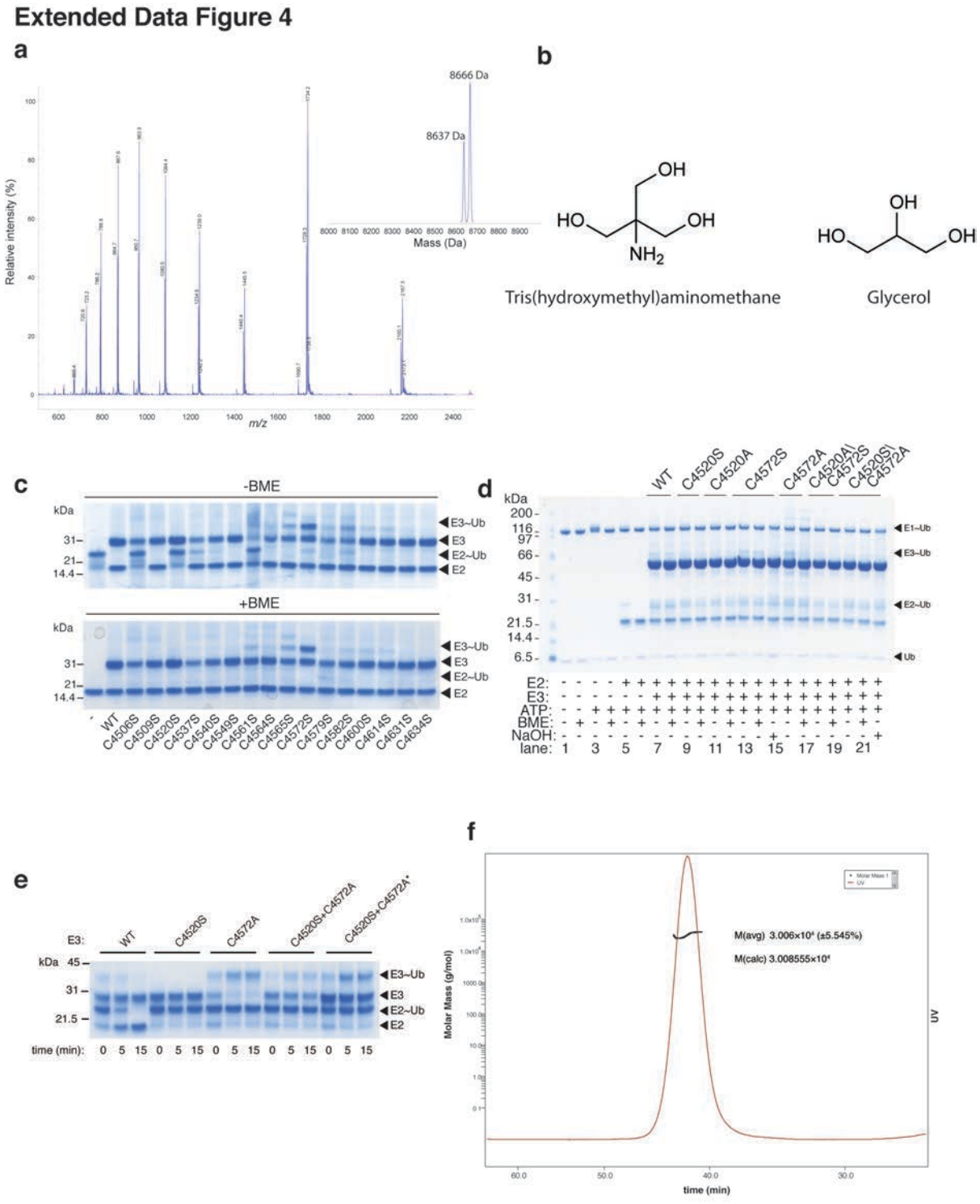
c

	+	++	+++
b1	114.0913	57.5493	38.702
b2	277.1547	139.081	93.0564
b3	414.2136	207.6104	138.7427
b5	625.3093	313.1583	209.1079
b6	738.3933	369.7003	246.8026
b7	852.4363	426.7218	284.8169
b8	939.4683	470.2378	313.8276
b9	1053.5112	527.2592	351.8419
b10	1110.5327	555.77	370.8491
b11	1197.5647	599.286	399.8598
b12	1310.6488	655.828	437.5544
b17	2894.4037	1447.7055	965.4727
B3	2637.3025	1319.1549	879.7723
B4	2766.3451	1383.6762	922.7865
B6	2894.4037	1447.7055	965.4727
y1	175.119	88.0631	59.0445
y2	288.2030	144.6051	96.7392
y3	401.2871	201.1472	134.4339
y4	516.3140	258.6606	172.7762
y5	629.3981	315.2027	210.4709
y6	1758.8738	879.9406	586.9628
y7	1871.9579	936.4826	624.6575
y8	1958.9899	979.9986	653.6682
y9	2016.0114	1008.5093	672.6753
y10	2130.0543	1065.5308	710.6896
y11	2217.0863	1109.0468	739.7003
y12	2331.1293	1166.0683	777.7146
y13	2444.2133	1222.6103	815.4093
y14	2558.2563	1279.6318	853.4236
y15	2655.3090	1328.1582	885.7745
y16	2792.3679	1396.6876	931.4608
Y1	175.1190	88.0631	59.0445
Y2	232.1404	116.5738	78.0517
Y3	303.1775	152.0924	101.7307
Y4	432.2201	216.6137	144.7449
Y5	503.2572	252.1322	168.4239
Y6	618.2842	309.6457	206.7662



Extended Data Figure 3. ABPs label C4520 within MYCBP2$_{cat}$ with high selectivity. (a) Recombinant MYCBP2$_{cat}$ was profiled with His-tagged ABPs based on the E2s UBE2D2 **5** and UBE2D3 **6**. Experiment repeated twice with similar results. (**b**) Putative active site cysteines

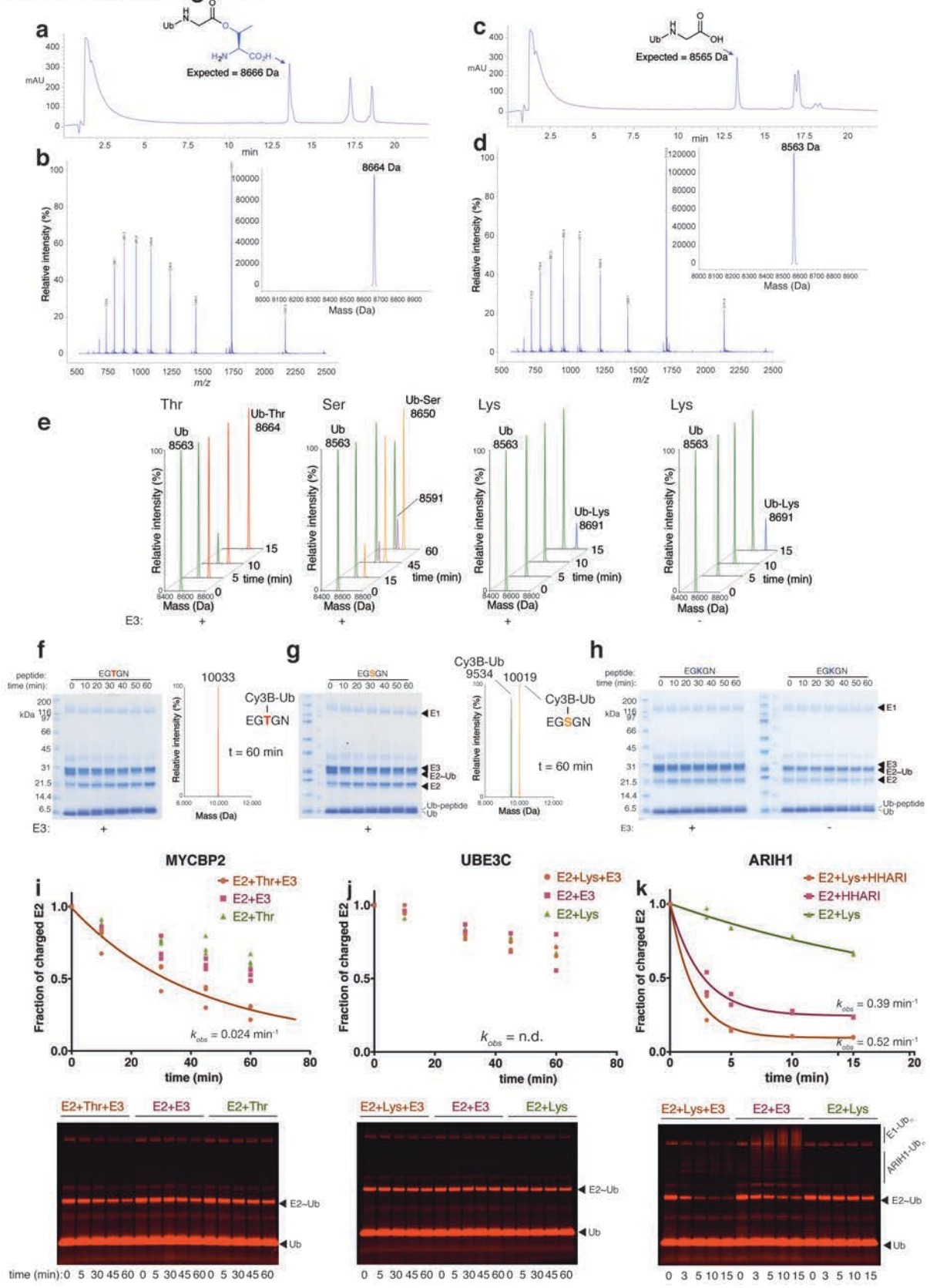
in MYCBP2 were determined by ABP profiling of a panel of cysteine to serine mutants. $MYCBP2_{cat}$ mutant (3 μM) was incubated with ABP **6** (12 μM) at 30 °C for 4 hours. ABP-treated samples were resolved by SDS-PAGE and visualized by Coomassie staining and immunoblotting against the hexahistidine reporter tag on the ABP. Mutation of 3 cysteine residues (C4506, C4520 and C4537) abolished ABP labelling. Asterisk corresponds to inadvertent cleavage of the hexahistidine tag from the ABP due to trace protease contamination of the E3 preparations. Experiment repeated twice with similar results. **(c)** Using the pLink software, 38 spectral matches corresponding to cysteine labelling sites in WT $MYCBP2_{cat}$ were identified. 36 of these corresponded to C4520. For the remaining two matches one corresponded to C4440, a predicted Zn-coordinating residue in the RING domain, and C4600, which did not significantly affect ABP labelled when mutated. The table lists the predicted and found fragment ions for the representative spectrum depicted in Figure 2b. The spectrum is for a 5^+ precursor ion (observed *m/z* = 614.5088; expected *m/z* = 614.5094). A mass tolerance of 20 ppm was applied for fragment ion assignment.



Extended Data Figure 4. Esterification activity of MYCBP2$_{cat}$ and further data in support of a dual cysteine mechanism that operates in *cis*. (a) Mass spectrum of condensation products between Ub and glycerol and, Ub and Tris. Expected mass for Ub condensation with glycerol is 8639 Da (found = 8637 Da). Expected mass for Ub condensation with Tris is

8668 Da (found = 8666 Da). Repeated twice with similar results. **(b)** Chemical structure of tris(hydroxymethyl)aminomethane and glycerol. **(c)** Discharge activity towards Tris/glycerol for all of the tested MYCBP2$_{cat}$ cysteine to serine mutants (selected mutants shown in Figure 2c). The C4506S mutation abolishes discharge activity but because C4506S resides within a Cys-X-X-Cys Zn-binding motif, this was assumed to be a structural rather than a catalytic defect. The C4561S mutation undergoes aberrant thioester adduct formation. An explanation is that C4561S mutation (also in a structurally important Cys-X-X-Cys Zn-binding motif) unfolds the protein and liberates Cys residues, that would otherwise be occupied as Zn ligands, that undergo aberrant transthiolation. Repeated twice with similar results. **(d)** Coomassie stain of the thioester/ester trapping assay with GST-MYCBP2$_{cat}$. After in-gel fluorescence scan, as shown in Figure 2e, the gel was Coomassie stained. Repeated >3 times with similar results. **(e)** The RCR E3 ligase activity is dependent on both C4520 and C4572. Combination of an inactive C4520S mutant with an inactive C4572A mutant did not restore activity. Hence there appears to be *cis*-cooperation between these two residues (*corresponds to elevated concentrations of E3 mutants). Repeated twice with similar results. **(f)** In further support of *cis*-cooperation, Size Exclusion – Multi Angle Light Scattering (SEC-MALS) data for untagged MYCBP2$_{cat}$ was consistent with a monodisperse species with a calculated molecular weight of 30.06 ($\pm$6) kDa (theoretical molecular weight of MYCBP2$_{cat}$ monomer = 30.08 kDa). Repeated twice with similar results.

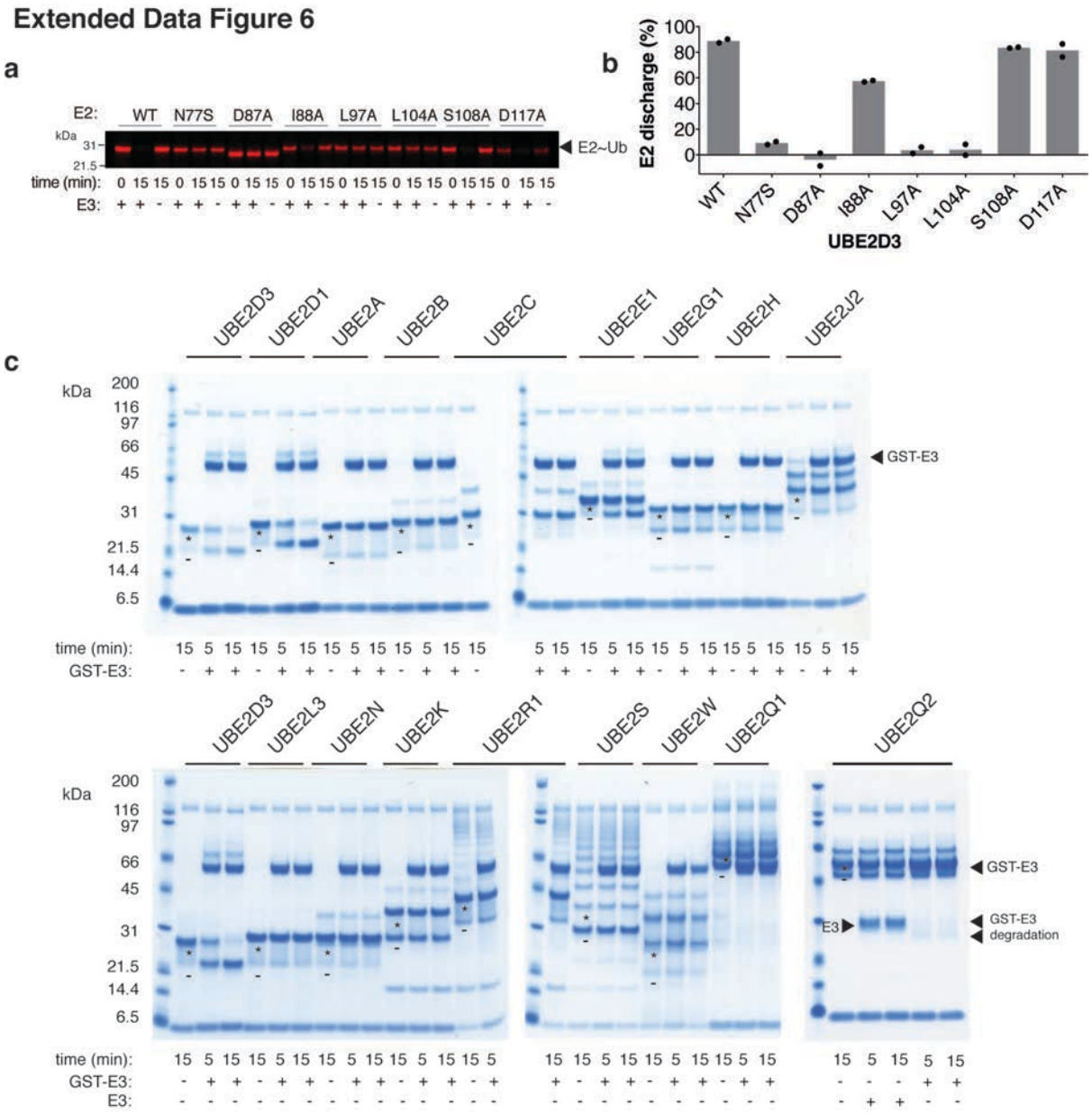
Extended Data Figure 5



Extended Data Figure 5. MYCBP2 has serine/threonine Ub esterification activity but has preference for threonine. (a) HPLC chromatogram of MYCBP2$_{cat}$ WT discharge reaction onto

threonine (50 mM). N.B. esterified threonine with a free amino terminus can undergo O-N acyl transfer forming a peptide-linked species. (**b**) Integrated single quadrupole electrospray ionization mass spectrum of the entire peak highlighted in the above chromatogram. Inset is the deconvoluted mass spectrum (as shown in **Figure 3b**) Expected mass of ubiquitinated threonine = 8666 Da; found mass = 8664 Da. (**c**) HPLC chromatogram of $MYCBP2_{cat}$ C4520S discharge reaction in the presence of threonine (50 mM). (**d**) Integrated single quadrupole electrospray ionization mass spectrum of the entire peak highlighted in the above chromatogram. Inset is the deconvoluted mass spectrum (as shown in Figure 3b). Expected mass of unmodified ubiquitin= 8565 Da; found mass = 8563 Da. All of the above experiments repeated 3 times with similar results. **(e)** Deconvoluted mass spectra for Ub species in the presence of amino acid (50 mM). The intensities of the Ub reactant and product are reflective of their relative abundance. Ubiquitinated serine (Ub-Ser) observed molecular weight = 8650 Da; theoretical = 8652 Da. Observed mass at 8591 Da corresponds to a side product that is only observed after extended incubation. Ubiquitinated lysine (Ub-Lys) observed molecular weight = 8691 Da; theoretical = 8693 Da. Assuming exponential Ub consumption, $t_{1/2}$ is ~5 min for threonine. For serine, $t_{1/2}$ is 10-fold slower. Lysine ubiquitination is E3-independent as a similar degree of modification is observed in the absence of E3. Repeated twice with similar results. **(f)** Coomassie stain of threonine gel presented in **Figure 3c**. Accompanying is the deconvoluted mass spectrum representative of all Ub species at the 60 minute time point. Observed mass of Cy3B-Ub modified threonine peptide = 10033 Da; theoretical mass = 10036 Da. **(g)** Coomassie stain of serine gel presented in Figure 3c. Accompanying is the deconvoluted mass spectrum representative of all Ub species at the 60 minute time point. Observed mass of Cy3B-Ub = 9534 Da; theoretical mass = 9537 Da.Observed mass of Cy3B-Ub modified serine peptide = 10019 Da;

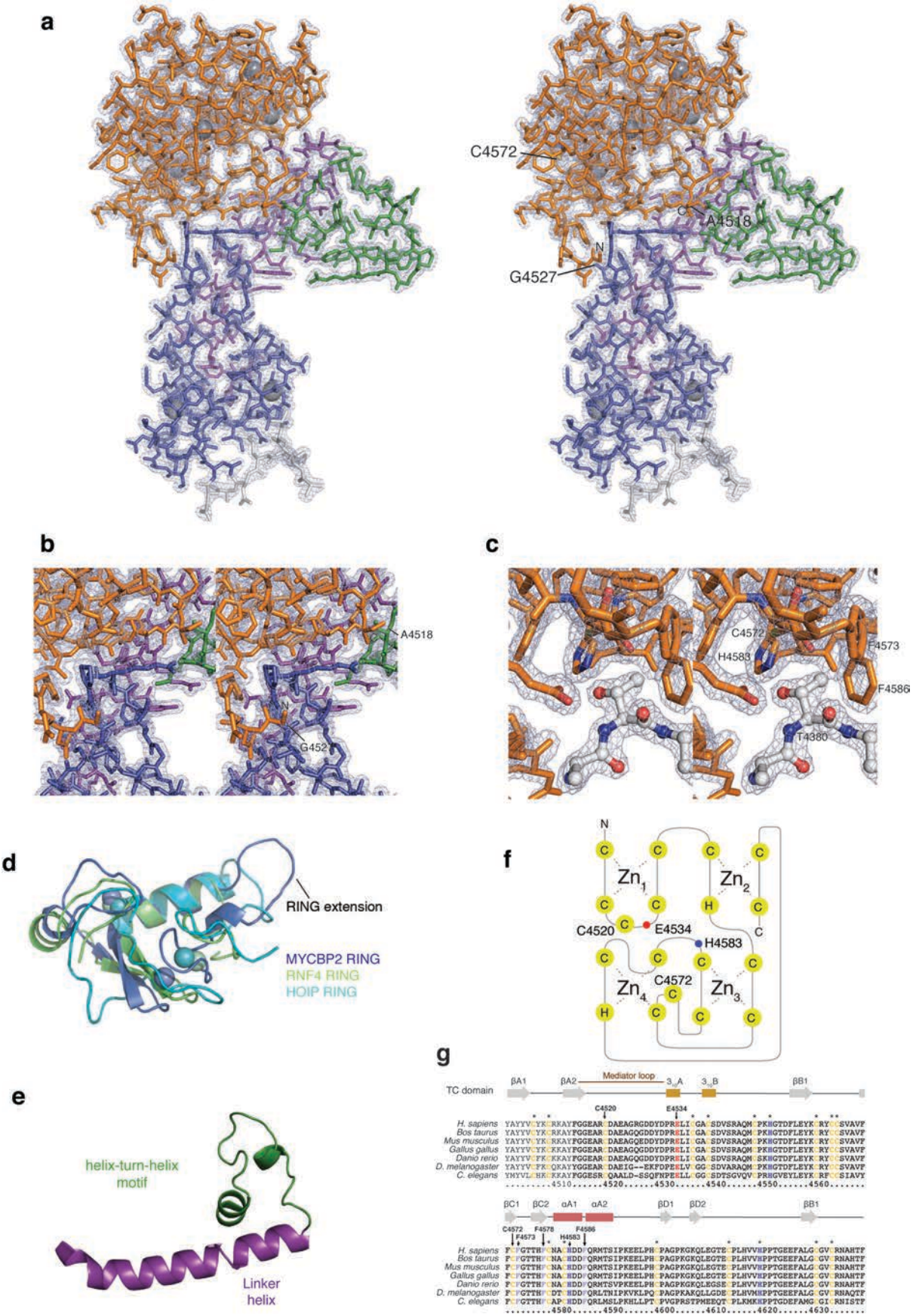
theoretical mass = 10022 Da. **(h)** Coomassie stain of lysine gels, in the presence and absence of E3, presented in Figure 3c. Inefficient modification of the lysine peptide is observed which is moderately enhanced in the absence of E3. Experiments f-h repeated > 3 times. **(i)** Top, observed rate constant (0.024 min^{-1}) for $MYCBP2_{cat}$-threonine mediated single turnover E2~Ub discharge determined by in-gel fluorescence of Cy3b labelled Ub. E2 was UBE2D3 and threonine (50 mM) was the substrate, (n = 3). Below, representative replicate gel used for quantification. **(j)** Top, observed rate constant for UBE3C-lysine mediated single turnover E2~Ub discharge was too slow to measure. E2 was UBE2L3 and lysine (50 mM) was the substrate, (n = 2). Below, representative replicate gel used for quantification. **(k)** Top, observed rate constant (0.52 min^{-1}) for HHARI-lysine mediated single turnover E2~Ub discharge. E2 was UBE2L3 and lysine (50 mM) was the substrate, (n = 2). The major component of this rate is attributable to autoubiquitination of lysine residues within HHARI because when lysine is withheld, k_{obs} HHARI mediated E2~Ub discharge is 0.39 min^{-1} and this is only partially outcompeted by the addition of lysine (n corresponds to the number of biologically independent experiments).



Extended Data Figure 6. E2 requirements of MYCBP2. (a) To establish whether RCR E3 ligase activity occurs exclusively via the proposed E3~Ub thioester intermediates, or alternatively mediates direct transfer of Ub from E2~Ub, the latter being a characteristic of RING E3s, we tested RCR E3 activity with a number of UBE2D3 mutants (N77S, D87A, I88A, L97A, L104A, S108A and D117A) that can be diagnostic for these two scenarios[5]. Single turnover E2~Ub discharge assays employing Cy3b-labelled Ub demonstrate that MYCBP2 has E2 requirements that are neither consistent with a HECT/RBR nor a RING mechanism. The N77 and D117 mutations alter E2 amino acids involved in pK_a suppression of the acceptor nucleophile and

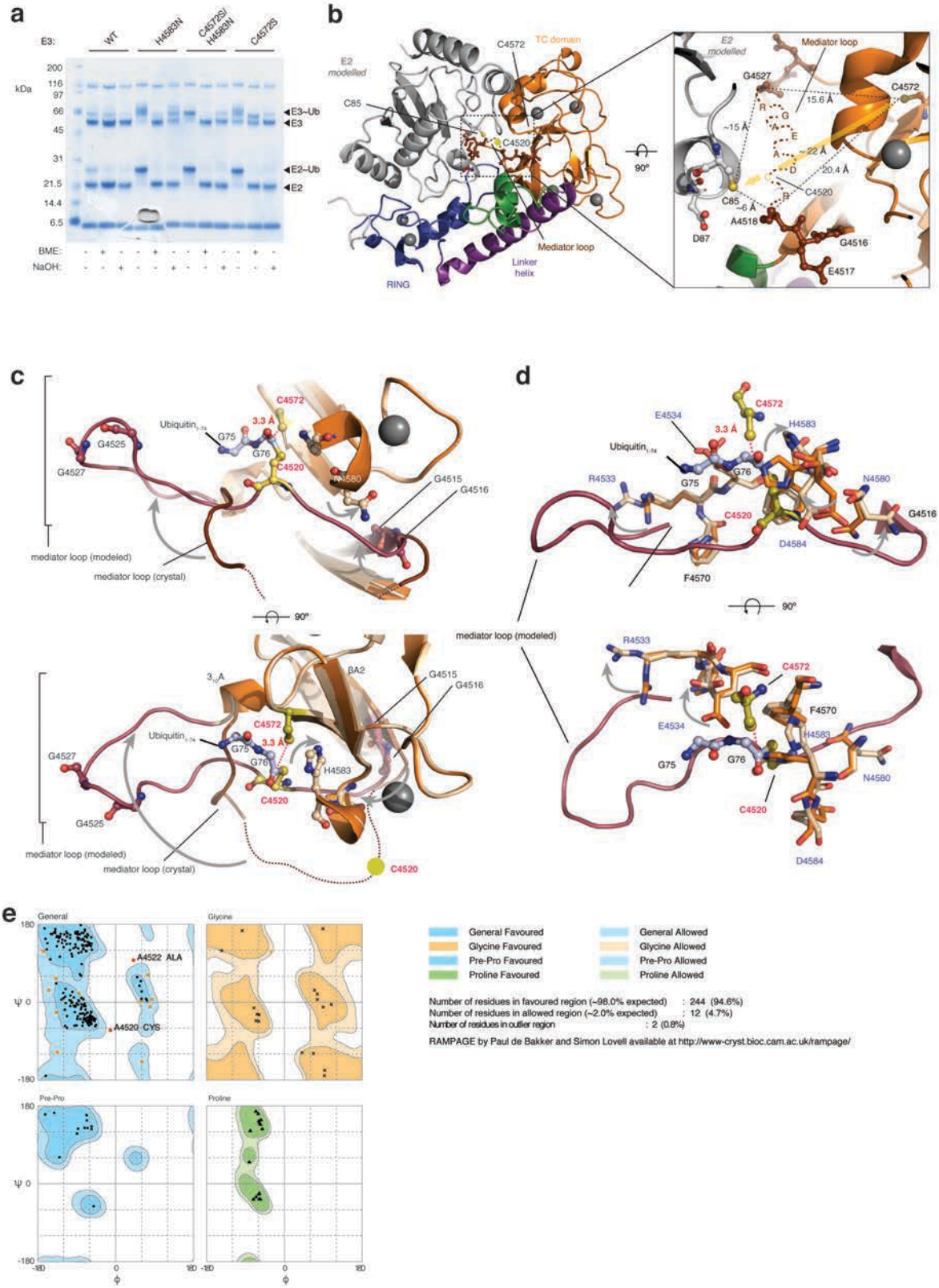
are required for RING activty[35,36]. Additionally, a characteristic of RING activity is the adoption of a “closed” E2~Ub conformation that involves E2 residues D87, I88, L97, L104 and S108. Unlike the RING requirements, UBE2D3, S108 and D117, were dispensable for E2-E3 transthiolation activity whereas the I88A mutant had reduced activity, the N77S mutant had strongly impaired activity, whereas the D87A, L97A and L104A mutants had negligible activity. Furthermore, RBR E3 activity is permissive to the L104A mutant[18]. Thus, based on our current understanding of the E2 requirements of these E3 classes, MYCBP2 has E2 requirements that are neither consistent with a HECT/RBR-like mechanism nor a RING-like mechanism. However, we cannot formally exclude that MYCBP2cat induces a closed E2~Ub conformation, characteristic of RING E3s, as it does not contain a prohibitive RING domain loop insertion[19]. **(b)** Quantification of the different E2 mutant activities. Mean of % E2~Ub discharge, (n = 2 biologically independent experiments). **(c)** . A total of 17 E2s were tested for threonine discharge activity with GST-MYCBP2$_{cat}$. UBE2D1, UBE2D3 and UBE2E1 were the only E2s that demonstrated detectable activity. (- corresponds to unmodified E2 and *corresponds to ubiquitin-charged E2). Unexpectedly, the HECT/RBR specific E2, UBE2L3, had negligible activity with MYCBP2. Certain E3s undergo E3-independent polyubiquitin chain formation and/or autoubiquitination. In the presence of UBE2Q2, GST-MYCBP2 undergoes minor degradation resulting in the appearance of two lower molecular weight species. Consequently, we also carried out the assay with untagged MYCBP2 which did not undergo degradation. The result that UBE2Q2 does not support MYCBP2 activity was the same. Experiment repeated twice with similar results.

Extended Data Figure 7



Extended Data Figure 7. Structural comparison and representative stereo views of the MYCBP2$_{cat}$ crystallographic model. (a) Wide field view. The distinct regions are colored in stick representation: RING domain (blue), linker helix (purple), helix-turn-helix motif (green) and tandem cysteine (TC) domain (orange). The mesh represents the experimental $2|F_{obs}|-|F_{calc}|$ electron density map contoured at 1.5 σ. C4572 is the downstream active cysteine residue in the esterification site. The mediator loop region is formed between A4518 and G4527 and is disordered in the structure. (**b**) Close up of the mediator loop region. (**c**) Close up of the esterification site. Threonine 4380 motif from the symmetry related molecule (T4380$_{sym}$) is shown and represented in grey ball and stick. **(d)** Superposition of MYCBP2 RING domain with the RING domain from the canonical RING E3 ligase RNF4[16], and from the RBR E3 ligase HOIP[18]. **(e)** The linker helix and helix-turn-helix motif that connect the RING domain to the Tandem Cysteine (TC) domain. **(f)** Diagram depicting Zn coordination network for the TC domain. Catalytic residues (numbered) are distributed throughout the TC polypeptide. **(g)** The Tandem Cysteine (TC) domain that confers threonine specificity is present in all MYCBP2 orthologues. All residues shown to be required for threonine esterification activity are conserved. Asterisks correspond to Zn-binding residues, grey arrows correspond to β-strands, gold rectangles correspond to 3_{10}-helices, and red cylinder corresponds to α-helix.

Extended Data Figure 8



Extended Data Figure 8. Modelling of E2-MYCBP2$_{cat}$ complex. (a) Ub adduct formation for catalytic mutants of GST-tagged MYCBP2$_{cat}$. The H4583N mutant undergoes near

quantitative Ub adduct formation. The adduct is largely removed after thiol treatment indicating that Ub is thioester-linked to the E3. The diffuse nature of the upper band might be due to the presence of a trapped thioester-linked Ub on C4520, and C4572, as the H4583N mutation prevents substrate deprotonation. The C4572S/H4583N mutant only forms a single Ub adduct that is thioester-linked, presumably to the C4520 residue. This illustrates that formation of the engineered ester-linked adduct on a mutated S4572 residue is dependent on the presence of a general base. C4520 does not appear to have a base in proximity hence its activity could be due to its intrinsic pK_a which results in it being nucleophilic at physiological pH in the absence of a general base. This could explain why we fail to form an engineered ester adduct on a mutant S4520 residue as serine is fully protonated at physiological pH. Experiment repeated twice with similar results. **(b)** Superposition of the RING domain from the RBR E3 ligase HOIP in complex with E2 (Ub linked to E2 has been omitted due to direct clash with the TC domain[18]; PDB ID: 5EDV) allows modelling of the E2 into our structure (grey cartoon representation). The catalytic C85 residue in E2 (mutated *in silico* from Lys to Cys[18]) is proximal to C4520 which undergoes transthiolation with E2~Ub. The right-hand panel is a top-down close up of the mediator loop region. The eight missing residues that form the mediator loop are schematically shown in brown text. **(c)** Model of the proposed Ub relay intermediate as shown in **Figure 5e** but from an additional perspective. For the experimental structure, TC domain residues are in orange and mediator loop residues are in dark brown. For the model, TC domain residues are in light orange and mediator loop residues are in mauve. The modeled E2, based on the superposition in (d), is in grey cartoon. Essential cysteines, C4520 and downstream C4572, are in yellow and colored by atom type. Ub residues (G75-G76) are in blue ball and stick notation and are colored by atom type. Gly residues in the mediator loop

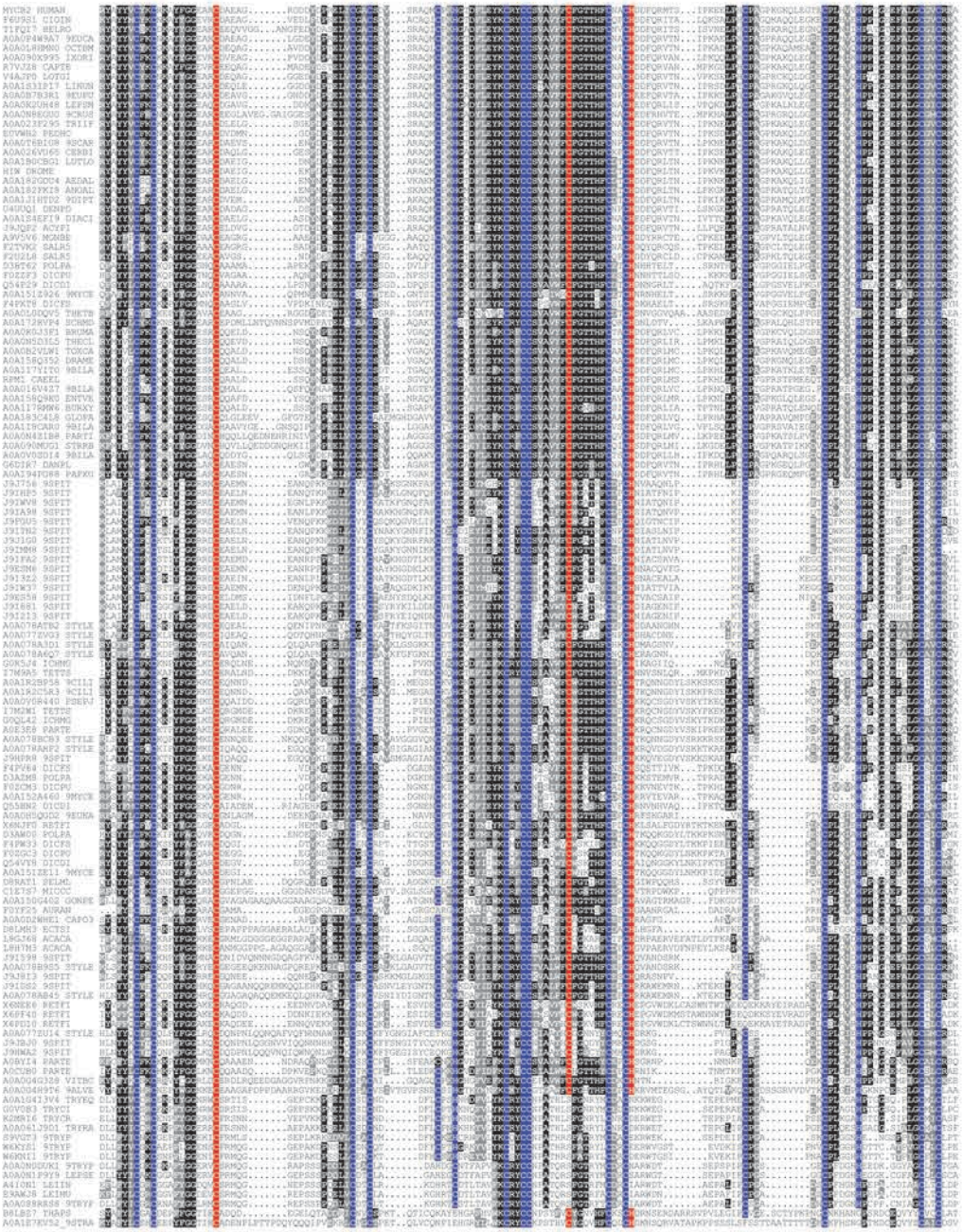
that are likely to be important for loop mobility are displayed in mauve ball and stick and colored by atom type. N4570 and H4583 side chains have been rotated by the specified angles to relieve steric clash. **(d)** As panel **c** but amino acid side chains that have been flipped to relieve steric clash with the modelled mediator loop are labelled in blue. (**e**) All phi and psi angles in the modelled structure fall within accepted values as determined by Ramachandran analysis with the RAMPAGE server[41].

Extended Data Table 1

	$MYCBP2_{4378-4640}$
Data collection	
Space group	$P6_1$
Cell dimensions	
a, b, c (Å)	82.58, 82.58, 103.30
α, γ, β (°)	90.00, 90.00, 120.00
Resolution (Å)	40-1.75 (1.78-1.75)
R_{merge}	0.051 (0.404)
I / σI	18.3 (2.4)
Completeness (%)	99.3 (93.3)
Redundancy	6.9 (3.0)
Refinement	
Resolution (Å)	40.0-1.75
No. reflections	40007
R_{work} / R_{free}	0.172/0.196
No. atoms	
Protein	1988
Ligand/ion	6
Water	257
B-factors	
Protein	28.2
Ligand/ion	26.6
Water	35.5
R.m.s. deviations	
Bond lengths (Å)	0.019
Bond angles (°)	1.9

Extended Table 1. Data collection and refinement statistics. Data were collected from a single crystal. Values in parentheses are for highest-resolution shell.

Extended Data Table 2



Extended Table 2. Structural alignments of MYCBP2 TC domain. Representative tandem cysteine domain sequences (130) from various taxa were isolated and aligned as described above. In virtually all animals, a single MYCBP2 orthologue is the only TC domain-containing gene. There is a relative large subfamily of sequences found in ciliates and a few other protists. Typically, each of these organisms contains multiple genes of this family. These proteins are also shorter than the animal versions. The alignment was rendered by Boxshade. Residues 50 % of invariant or conservatively substituted are on black and grey background, respectively. Finally, Cys/His residues involved in Zn coordination are indicated in blue, while the Cys/His residues involved in the ubiquitin relay are shown in red.